GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 13:25:53 ; Search time 3903 Seconds
                                             (without alignments)
                                             9663.111 Million cell updates/sec

Title: US-09-913-772-1
Perfect score: 1035
Sequence: 1 atgaaagcaatttcgtact........taactcagccggcgggtaa 1035

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_htc:*
                 4: gb_est3:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		461.6	44.6	671	7	CF558062	CF558062 1115033F0
2		454.8	43.9	628	7	CF565551	CF565551 1115101B0
3		426	41.2	639	7	CF565552	CF565552 1115101B0
4		396	38.3	591	7	CF558063	CF558063 1115033F0
5		384.2	37.1	767	9	CL656958	CL656958 PRI0128a
6		274	26.5	422	7	CF561913	CF561913 1115069D0
7		265.4	25.6	438	7	CF563205	CF563205 1115082A0
8		219	21.2	730	9	CL657503	CL657503 PRI0129d
9		176	17.0	323	7	CF563206	CF563206 1115082A0
10		159	15.4	603	2	BE423296	BE423296 WHE0063_E
11		134.2	13.0	655	8	AF075794	AF075794 AF075794
c 12		61.4	5.9	601	7	CF915072	CF915072 B0973F11-
c 13		61.4	5.9	601	7	CF915603	CF915603 B0982E05-
14		61.4	5.9	736	7	CF544559	CF544559 Hd_mxl7_6
15		61.4	5.9	763	7	CK326686	CK326686 Hd_mx24_0
16		61.4	5.9	860	9	CN824451	CN824451 Oa_splbn
c 17		56.6	5.5	1047	8	BZ561117	BZ561117 pacs2-164
18		54.8	5.3	748	8	BZ570213	BZ570213 msh2_1236
c 19		54.8	5.3	1030	8	BZ554003	BZ554003 pacs1-60
c 20		53.4	5.2	773	8	BZ562758	BZ562758 pacs2-164
21		53.2	5.1	912	8	BZ570029	BZ570029 msh2_1132
22		51.6	5.0	134	9	CR337045	CR337045 Medicago
c 23		51.6	5.0	1772	8	BZ575847	BZ575847 msh2_4654
c 24		49.6	4.8	600	7	CK991504	CK991504 EST0056_E
c 25		48.6	4.7	651	8	BZ566271	BZ566271 pacs2-164
c 26		48.4	4.7	1200	8	BZ558860	BZ558860 pacs2-164
c 27		47.4	4.6	296	1	AI904113	AI904113 IL-BT042-
c 28		46.8	4.5	840	9	CL666637	CL666637 PRI015a
29		44.8	4.3	1264	8	BZ569964	BZ569964 msh2_1103
30		43.8	4.2	1232	8	BZ555345	BZ555345 pacs1-60
c 31		42.2	4.1	611	6	CB853355	CB853355 UI-CF-FN0
c 32		42.2	4.1	730	7	CF891479	CF891479 UI-CF-FN0
c 33		42	4.1	910	7	BZ673049	BZ673049 PUBEH3lTD
c 34		41.8	4.0	467	7	CK453782	CK453782 911024_MA
35		41.8	4.0	775	8	BZ575815	BZ575815 msh2_464.
36		41.4	4.0	1118	8	BZ567231	BZ567231 pacs2-164
37		40.4	3.9	1144	8	BZ552769	BZ552769 pacs1-60
38		40.2	3.9	774	8	BZ576217	BZ576217 msh2_483.
39		40.2	3.9	1146	8	BZ569714	BZ569714 pacs2-164
40		40.2	3.9	1216	8	BZ569713	BZ569713 pacs2-164
41		39.8	3.8	794	8	BZ576305	BZ576305 msh2_487.
c 42		39.8	3.8	1040	8	BZ560041	BZ560041 pacs2-164
43		39.6	3.8	976	8	BZ552409	BZ552409 pacs1-60
44		39.6	3.8	1121	5	BQ436127	BQ436127 AGENCOURT
c 45		39.4	3.8	533	4	BJ192697	BJ192697 BJ192697

ALIGNMENTS

RESULT 1
CF558062

LOCUS   CF558062   671 bp   mRNA   linear   EST 23-SEP-2003
DEFINITION 1115033F05.x1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA
            sequence.
ACCESSION   CF558062
VERSION   CF558062.1 GI:34984145
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
  ORGANISM Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 671)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Merchant,S., Quinn,J. and Shrager,J.
  TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1115
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 671.
FEATURES   Location/Qualifiers
     source   1..671
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant(1998) Methods in
                     Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
                     micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
                     Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,

```
Db   560 CCCTGAAACCGGAAGTCAGGCTGCTGCTCTGGATCAGCTGTACAGCCAGCTGAGCAACCTGG 619
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Qy   740 ATCCGAAAGACGGTTCCGCTTGTTGTTCTCGGGCTACACCGACCGATCGGTTC 791
         |||||||||||||||||| |||||||||||||| |||||||||||  |||||||
Db   620 ATCCGAAAGACGGTTCCGTAGTTGTTCTCGGGTTACACCGACCACCGGATCGGTTC 671
```

RESULT 2
CF565551
Qy   CF565551
LOCUS   1115101B09.x1 C. reinhardtii CC-1690 (mt+), Stress III   628 bp   mRNA   linear   EST   23-SEP-2003
DEFINITION 1115101B09.x1 C. reinhardtii CC-1690 (mt+), Stress III (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA sequence.
ACCESSION   CF565551
VERSION   CF565551.1   GI:34991634
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
ORGANISM   Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 628)
AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Merchant,S., Quinn,J. and Shrager,J.
TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1115
JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 628.
FEATURES   Location/Qualifiers
   source   1..628
            /organism="Chlamydomonas reinhardtii"
            /mol_type="mRNA"
            /strain="CC-1690 wild type mt+ 21gr"
            /db_xref="taxon:3055"
            /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III (normalized), Lambda Zap II"
            /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition III library, constructed by Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from CC-1690 cells grown to mid-log phase in copper-free TAP medium (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated (about 100 micromole/m2/sec) incubator at 22 C (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279);
            CC-1690 cells grown to mid-log phase in low Fe (1 micromolar Fe) TAP medium (see La Fontaine S, Quinn JM, Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J, Merchant S.LaFontaine et al. (2002) Eukaryotic Cell, 1:736-757) in a shaking illuminated incubator (same conditions as above). CC-1690 cells were grown to mid-log phase in TAP medium in a shaking illuminated incubator to a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089); CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high
```

Merchant S.LaFontaine et al. (2002) Eukaryotic Cell, 1:736-757) in a shaking illuminated incubator (same conditions as above). CC-1690 cells were grown to mid-log phase in TAP medium in a shaking illuminated incubator to a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089); CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified from each sample, pooled and cDNA synthesized (see Shrager et al, Plant Physiol. 131, 401-408 for details). The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

```
Query Match          44.6%;    Score 461.6;    DB 7;    Length 671;
Best Local Similarity 83.3%;   Pred. No. 5.7e-124;
Matches    543;    Conservative    0;    Mismatches    94;    Indels    15;    Gaps    1;

Qy   140 ATCAGCTTGGTGCTGCTGCGTTCGGTGGTTACCAGTTAACCCGTACCCGGTTTCGAAA 199
         ||||||||||||| |||||||||   ||||||||||||||||||||||   ||||||||
Db    35 ACCAACTGGGCGGGCCTGCGTGCTTTTGGTGGTTACCAGTTAACCCGTATGTGGCTTTGAAA 94

Qy   200 TGGGTTATGACTGGCTGGGGCCGTATGGCGTAAAGGCAGCGTTGACAACGGTGCTTTCA 259
         ||||||| ||||||| || | |||||| |||| ||||||||||||| ||||||||  | |
Db    95 TGGGTTACGACTGGTTAGGTCGTATGCCGTACAAAGGCAGCGTTGAAAACGGTGCATACA 154

Qy   260 AAGCTCAGGGCGTTCAGCTGACACCGCTAAACTGGTTACCCGGTACCCGATCACTGGACA 319
         |||||| ||||||||| ||||||||||||||||| ||||||  ||||| ||||| |||||
Db   155 AAGCTCCAGGGCGTTCAACTGACACCGCTAAACTGGGTTACCCAATCACTGACGACCGGACA 214

Qy   320 TCTACACCCGTCTGGGCGGCATGGTTGGCGCGCTGACTCCAAAGGCAACTACGGTTCTA 379
         ||||||| |||||||||||||||||| ||||  |||| |  ||| |||   |  ||
Db   215 TCTACACTCGTCTGGGCGGCATGGTAGGCGTGCAGACACTAAATCCAA----------- 263

Qy   380 CCGGGCGTTTCCGTAGGCAACACGACACCGGCGCGGTTCCCAGTATTGCTGGCGGCGTAG 439
         ||| ||||| ||  | |||  ||||||||||||   ||   | | | |||||||| |||
Db   264 ---CGTTTATGGTAAAACCACGACACCGGCGCTCTTCGCTGGCGGTGTTG 319

Qy   440 AGTGGGCGTGTTACTCCGTGACATCCTGGAATACCAGTGGGTTAACACATCG 499
         ||| |||| |  || |||| ||||| || ||||||||||| | ||| |||||
Db   320 AGTACGGCGGATCACTCCTGAAATCGCTCCGTCTGGAATACCAGTGGGACCAACATCG 379

Qy   500 GCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATGCTGGAGCCTGGGCGTTTCCT 559
         | |||||||||||| | | | |||| | | |||  || |||| ||||||||||  |||||
Db   380 GTGACCGCCACACCATCGGCACTCTCGTCCGGACAAACCGCATGCTGCTGGAGCCTGGGTGTTTCCT 439

Qy   560 ACCGGCTTCGGTCAGGAAGATGCTGCACCGGCGTTGTTGCTCCGGCTCCGGCTCCGG 619
         |||| ||||||||||| | || || || ||| |  |||||||| |||| | |||| |
Db   440 ACCGTTTCGGTCAGGGCGAAGCAGCTCCAGTTGCTCCGGCTCCAGCTCCGGCACCGG 499

Qy   620 AAGTTGGCTACCAAGCACTTCACCCTGAAGTCGACGTTCGTTCAACTTCAACAAAGCTA 679
         ||||  | ||| |||||||||||| |||||||| | ||||| ||||||||||||||| |
Db   500 AAGTACAGACAAGCACTTCACCTGAAGTCGAAGTCGACTTCGTTCAACTTCAACAAAGCAA 559

Qy   680 CCCTGAAACCGGAAGTCAGGCTGCTGCTCTGGATCAGCTGTACACTGAGCAACATGG 739
         |||||||||||||||| ||
Db   680
```

REFERENCE
AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
           Merchant,S., Quinn,J. and Shrager,J.
TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model,
           Unicellular System for Analyzing Gene Function and Regulation in
           Vascular Plants. Project: 1115
JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
           DCMB Box 91000
           Duke University
           Durham, NC 27708-1000
           Tel: 919 613 8159
           Fax: 919 613 8177
           Email: chauser@duke.edu
           High quality sequence stop: 639.

FEATURES   Location/Qualifiers
     source   1..639
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant(1998) Methods in
                     Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
                     micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
                     Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,
                     Merchant S.LaFontaine et al. (2002) Eukaryotic Cell,
                     1:736-757) in a shaking illuminated incubator (same
                     conditions as above). CC-1690 cells were grown to mid-log
                     phase in TAP medium in a shaking illuminated incubator to
                     a density of 8x10e6 cells/ml. The flask was transferred to
                     a shaking platform (200 rpm) at room temperature (23C) 12
                     micromole/m2/sec illumination and bubbled in a stoppered
                     flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour
                     prior to harvesting for RNA isolation (as per Quinn JM,
                     Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC
                     275:6080-6089); CC-1690 cells grown to mid-log phase
                     (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x10e6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to
                     mid-log phase in HS medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x10e6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified
                     from each sample, pooled and cDNA synthesized (see Shrager
                     et al, Plant Physiol. 131, 401-408 for details). The cDNA
                     was directionally cloned into lambda Zap II (Stratagene)
                     in the EcoRI (5') and XhoRI (3') sites.pBluescript II SK-
                     plasmids were excised from the lambda ZAP clones by
                     superinfection with ExAssist (Stratagene) phage. The
                     library was normalized using method 4 described in Bonaldo
                     et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match   41.2%; Score 426; DB 7; Length 639;
Best Local Similarity   86.3%; Pred. No. 1.6e-113;
Matches   471; Conservative   0; Mismatches 75; Indels   0; Gaps   0;

Qy   490 AACAACATCGGGCACGCGGGCACTGTGGTACCCGTCCTGATAACGGCATGCTGAGCCTG 549

Db   638 AACAACATCGGTGACGACGACCACCACCATCGGACTCGTCCGGACACACGGCATGCTGAGCCTG 579

---

light (11000 mol photon/m2/s) with shaking (150 rpm) and
sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified
from each sample, pooled and cDNA synthesized (see Shrager
et al, Plant Physiol. 131, 401-408 for details). The cDNA
was directionally cloned into lambda Zap II (Stratagene)
in the EcoRI (5') and XhoRI (3') sites.pBluescript II SK-
plasmids were excised from the lambda ZAP clones by
superinfection with ExAssist (Stratagene) phage. The
library was normalized using method 4 described in Bonaldo
et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match   43.9%; Score 454.8; DB 7; Length 628;
Best Local Similarity   83.3%; Pred. No. 5.5e-122;
Matches   535; Conservative   0; Mismatches 92; Indels 15; Gaps   1;

Qy   145 CTTGGTGCTGGTGCGTTCGGTGGTTACCCGTTAACCCGTACCTCGGTTTCGAAATGGGT 204

Db   2 CTGGGGCGCTGGTGCTTTTGGTGCGTTACCAGGTTAACCCGTATGTTGGCTTTGAAATGGGT 61

Qy   205 TATGACTGGCTGGCCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCT 264

Db   62 TACGACTGGTTAGGTCGTATGCCGTACAAAGGCAGCGTTGAAAACGGTGCATACAAAGCT 121

Qy   265 CAGGGCGGTTCAGCTGACCGGCTAAACTGGTTACCGGCGCTGACTCACGATCTGGACATCTAC 324

Db   122 CAGGGGCGGTTCAACTGACCGGCTAAACTGGTTACCCAATCACTCGACGACCTGGACATCTAC 181

Qy   325 ACCCGTCTGGGCGGCCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGGCTTCTACCCGGC 384

Db   182 ACTCGTCTGGGGTGCCATGGTATGGCGTGCAGACACTAAATCCAA---------C 226

Qy   385 GTTTCCCGTAGGAAACGACACGACCTGGCGTTCCCCAGTATTTGCTGCGGCGGGCTAGAGTGG 444

Db   227 GTTTATGTGTAAAAACCACGACACCGGCGGTTCTCCGGTTCTCCGGTGTTGAGTAC 286

Qy   445 GCTGTTACTCGTGACATCGCTACCCGTCCGGAATACCAGTGGTTAACACATCGGCGAC 504

Db   287 GCGATCACTCCTGAAATCCTACCCGTCCGGAATACCAGTGGTTAACACATCGGTGAC 346

Qy   505 GCGGGCACTGTGGTACCCGTCCGTCCGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC 564

Db   347 GCACACACCATCGCACTCGTCCGGACACGGCACTCGTGGGTGTTTCCTACCGT 406

Qy   565 TTCGGGTCAGGAAGGATGTGCACCGGTTGTTGCTCCGGCTCCGGCCTCCGGAAGTG 624

Db   407 TTCGGTCAGGGGCGGAAGCGCTCCAGTAGTTGCTTCCGGCTCCAGCTCCGGCACCGGAAGTA 466

Qy   625 GCTACCAAGCACTTCACCCTGAAGTCTGACGTTCGTTCAACTTCAACAAAGCTACCCTG 684

Db   467 CAGACCAAGCACTTCACTCTGAAGTCTGACGTTCGTTCAACTTCAACAAAGCAACCCTG 526

Qy   685 AAACCGGAAGGTCAGCAGGCTCTGACTGTGTACACTCAGCTGAGCAACATGGATCCG 744

Db   527 AAACCGGAAGG'TCAGGCTGTCTTGGATCAGGTGTACAGCCAGCCAGCGTGAGCAACCTGGATCCG 586

Qy   745 AAAGACCGGTTCCGCTGTTCCGTTGTTCTGGGCTACACCGACCGCCATC 786

-Db   587 AAAGACCGGTTCCGTAGTTGTTCTGGGGTTCTGGGCTACACCGACCGCCATC 628

RESULT 3
CF565552/c
LOCUS   CF565552   639 bp   mRNA   linear   EST 23-SEP-2003
DEFINITION 1151101B09.y1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 5', mRNA
            sequence.
ACCESSION   CF565552
VERSION   CF565552.1 GI:34991635
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
  ORGANISM Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.

(normalized), Lambda Zap II"
/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition III library, constructed by Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from CC-1690 cells grown to mid-log phase in copper-free TAP medium (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated (about 100 micromole/m2/sec) incubator at 22 C (see Quinn and Merchant(1998) Methods in Enzymology, 2997:263-279); CC-1690 cells grown to mid-log phase in low Fe (1 micromolar Fe) TAP medium (see La Fontaine S, Quinn JM, Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J, Merchant S.LaFontaine et al. (2002) Eukaryotic Cell, 1:736-757) in a shaking illuminated incubator (same conditions as above). CC-1690 cells were grown to mid-log phase in TAP medium in a shaking illuminated incubator to a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089); CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified from each sample, pooled and cDNA synthesized (see Shrager et al, Plant Physiol. 131, 401-408 for details). The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

```
ORIGIN

Query Match          38.3%;   Score 396;   DB 7;   Length 591;
Best Local Similarity 87.0%;   Pred. No. 9.8e-105;
Matches  435;   Conservative   0;   Mismatches  65;   Indels   0;   Gaps   0;

Qy  536 GCATGCTGAGCCTGGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTG 595
        |||||||| ||||| |||||||||||||| ||||||||||   |   |||||    |||||
Db  591 GCATGCTGATCCTGGGGTGTTCCTACCGCGTTCGGTCAGGGCGGCGAAGCAGCCCCAGTAGTTG 532

Qy  596 CTCCGGCTCCGGCTCCGGAAGTGGCTACCAGAGACCACTTCACCCTGAAGTCTGACG 655
        |||||||||| |||||||||||||| |||||| |||||||||| |||||||||||||
Db  531 CTCCGGCTCCAGCTCCGGCACCGGAGTACAGATACCACGCACTTCACTCTGAAGTCTGACG 472

Qy  656 TTCTGTTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGGCTCTGGATCAGC 715
        ||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db  471 TTCTGTTCAACTTCAACAAAGCAACCCTGAAACCGGAAGGTCAGGCTGCTCTGGATCAGC 412

Qy  716 TGTACACTCAGCTGAGCAACATGGATCCGGAAAGACGGTTCCGCTGTTGTTCTGGGCTACA 775
        ||||| ||||||||||||||  | |  ||| |||||||||||  ||||||||||||  |||
Db  411 TGTACAGCCAGCTGAGCAACCCTGAGCTTCCGGATCCGAAAGACGGTTCCGTAGTTGTTCTGGGGTTACA 352

Qy  776 CCGACCGCATGCGGTTCCGAGCTTACAACCAGCTGTCTGAGAAACGTGCTCAGTCCCG 835
        |||||||||||| |||||    |||||||||||  ||| ||  ||||||||| |||| |
Db  351 CCGACCGCATGCGGTTCTGACGGTTCACGCTTACAACCAGGGGTCTGTCCGAGCGCGTGCTGCAGTCTG 292

Qy  836 TTGTTGACTACCTGGTTGCTAAAGGCATCCGCGCTGCCAAAATCTCCGCTCCGGCATGG 895
        |||||||  |||||| | |||||||| ||| | || | ||||| |||||| |||| ||||
Db  291 TTGTTGATTACCTGATCTCCAAAGGTATCCCGGCACCAAGAATCTCCGCACGTGGTATGG 232

Qy  896 GTGAATCCAACCGCGTTACTGGCACACCTGTGACAACGTGAAAGCTGCGCCGCTGCCCTGA 955
        || |||||||||| |||||||||||||||||||||||  ||  |||| | | ||||| |||
```

RESULT 4
LOCUS   CF558063   591 bp   mRNA   linear   EST   23-SEP-2003
DEFINITION 115033F05.y1 C. reinhardtii CC-1690 (mt+), Stress III (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 5', mRNA sequence.
ACCESSION   CF558063
VERSION   CF558063.1 GI:34984146
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
ORGANISM   Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 591)
AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Merchant,S., Quinn,J. and Shrager,J.
TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1115
JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 591.
FEATURES   Location/Qualifiers
     source   1..591
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III

```
Qy  550 GGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGGCT 609
        ||| |||||||||| |||||||||||   |   |||||    |||||||||||
Db  578 GGTGTTCCTACCGCGTTCGGTCAGGGCGGCGAAGCAGCTCCAGTAGTTGCTCCGGCTCCAGCT 519

Qy  610 CCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCGTTGTTCAACTTC 669
        |||||||||||||||| ||| |    |||||| |||||||||||||  |||  ||||||||||
Db  518 CCGGCACCGGAGTACAGACCACTTCACTTCACTCTGAAGTCTGACGTTCGTTGTTCAACTTC 459

Qy  670 AACAAAGCTACCCTGAAACCGGAAGGTCAGGCTCTGGATCAGCTGTACACTCAGCTG 729
        ||||||| |||||||||||||||||||||||| ||||||||||||||| ||||||||
Db  458 AACAAAGCAACCCTGAAACCGGAAGGTCAGGCTGCTCTGGATCAGCCAGCCAGCTG 399

Qy  730 AGCAACATGGATCCGGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGACCGCATCGGT 789
        |||||| | |  ||| |||||||||||  ||||||||||||  |||||||||||||||||||
Db  398 AGCAACCCTGAGCTTCCGGATCCGAAAGACGGTTCCGTAGTTGTTCTGGGGTTACACCGACCGCATCGGT 339

Qy  790 TCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTG 849
        || ||  ||||||||||   ||  ||| ||  ||||||||| |||  |||||||  |||||
Db  338 TCTGACGCTTACAACCAGGGGTCTGTCCGAGCGCGTGCTGCAGTCTGTTGTTGATTACCTG 279

Qy  850 GTTGCTAAAGGCATCCGCGCTGGCAAAATCTCCGCTCCGGCATGGTGGTGAATCCAACCCG 909
        | || |||||| ||| | || | ||||| ||||| | | |||| |||| ||  | ||| ||
Db  278 ATCTCCAAAGGTATCCCGGCACAGATCTCCCACCGGTATGGGCAATCTAACCCG 219

Qy  910 GTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCCCTGATCGATTGCCTGGCT 969
        |||||||||||||||||||||||  || || | ||| |  ||||||| |||||||||
Db  218 GTTACTGGCAACACCTGTGACAACAGCGGTGCTGCACTGACTGCCTGGCT 159

Qy  970 CCGGATCGTCGTGTAGAGATCGAAGTCGAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGCG 1029
        |||||||||||| ||||||||||||| ||||| | |||||| |||||||||||||||| ||
Db  158 CCGGATCGTCGTCGTGCGTAGAGATCGAAGTTGAAGGTATCAAAGACGTTGTAACTCAGCCGCAG 99

Qy 1030 GGTTAA 1035
        | ||||
Db   98 GCTTAA 93
```

```
Db   231  GCGAATCCAACCCGGTTACTGGCAACACCTGTGACAACGCAGCGTGCTGCACTGA  172

Qy   956  TCGATTGCTGGCTCCGGATCGTCGTGTAGAGATCGAAGGTAAAGGCTACAAAGAAGTTG  1015
          ||||||| ||||||||||||||| | |||||||||||||| ||||| ||||||||| |||
Db   171  TCGACTGCTGGCTCCGGATCGTCGCGTAGAGATCGAAGGTTAAAGGTATCAAAGACGTTG  112

Qy  1016  TAACTCAGCGCGGCGGGGGTAA  1035
          ||||||||| |||| || ||||
Db   111  TAACTCAGCCGCGAGGGCTAA   92
```

RESULT 5
CL656958/c
LOCUS   CL656958   767 bp   DNA   linear   GSS 09-JUL-2004
DEFINITION PRI0128a_C05 - PRI0128a_B21 (767) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL656958
VERSION   CL656958.1   GI:50137970
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
ORGANISM   Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 767)
AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE   AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 004970716013 71
            Fax: 004970716014 98
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES   Location/Qualifiers
     source   1..767
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

Query Match   37.1%; Score 384.2; DB 9; Length 767;
Best Local Similarity 87.0%; Pred. No. 3.1e-101;
Matches 422;   Conservative 0;   Mismatches 63;   Indels 0;   Gaps 0;

```
Qy   551  GCGTTCCTACCGCTTCGGTCAGGAGATGCTGCACCGGTTGTTGCTCCGGCTCCGGCTC   610
          | ||| ||||||| ||||||||||  |  | |||| | |||||||||||||||| ||||
Db   767  GTGTTTCTACCGTTCGGTCAGGGCGAAGCAGCTCCAGTAGTTGCTCCGGCTCCAGCTC   708

Qy   611  CGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCA  670
          ||||| ||||  |  ||||| || |||||||| |||||||||||||||||||||||||||
Db   707  CGGCACCGGAGTGAGTACGACACCAAGCACTTCACTCTGAAGTCTGACGTTCTGTTCAACTTCA  648

Qy   671  ACAAAGGTACCCTGAAACCGGAAGGTCAGCAGGCTCGATCAGTCGTACACTCAGCTGA   730
          ||||| | |||||||||||||||||||||||||||||| || ||||||  |||||| |
Db   647  ACAAAGCAACCCTGAAACCGGAAGGTCAGCAGGCTCGGATCAGTCGTACAGCCAGCTGA   588

Qy   731  GCAACATGGATCCGAAAGACGGTTCCGCTGTCTTCCGGGGCTACACCGACCGCATCGGTT  790
          ||||| ||||||||||||||| ||||||  |  | |||| |||||||||||||||||||
Db   587  GCAACCTGGATCCGAAAGACACGGTTCCCTAGTTGTTCTGGGTTACACCGACCGCATCGGTT  528

Qy   791  CCGAAGCTTACAACCAGCAGTCGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGG  850
          | ||||||| |||||| |  |||   ||  |||||||| | ||| |||||||| |||||||
```

```
Db   527  CTGACGCTTACAACCAGGGTCGTCCGAGCGCCGCCGTCTCGTTGTTGATTACCTGA   468

Qy   851  TTGCTAAAGGCATCCGGCTGGCAAAATCTCCGCTCCGGCATGGGTGAATCCAACCCGG   910
          | |   |||| |||||||| | |||| || ||| | ||||| ||| |||||||||||
Db   467  TCTCCAAAGTATCCGGCAGACAAGATCTTCCGCACGTGTATGGGCGAATCCAACCCG   408

Qy   911  TTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCTGCCTGCCCGATCGATTGCCTGGCTC  970
          |||||||||||||||||||||||| |||| | | ||||| ||||||| ||||| ||||||
Db   407  TTACTGGCAACACCTGTGACAACAGCGTGAAACAGCAGGCTGCGTGCTGCACTGACTGCCTGCCTGGGCTC  348

Qy   971  CGGATCGTCGTGTAGAGATCGAAGGTTAAACGGCTACAAAGAAGTTGTAACTCAGCGCGGCGCG  1030
          |||||||||| ||||||||||||||||||||| |||||||| ||||||||||||||||  |||||
Db   347  CGGATCGTCGCGTAGAGATCGAAGGTTAAACGGTATCAAAGACGTTGTAACTCAGCCGCGCAGG   288

Qy  1031  GTTAA  1035
          | |||
Db   287  CTTAA   283
```

RESULT 6
CF561913
LOCUS   CF561913   422 bp   mRNA   linear   EST 23-SEP-2003
DEFINITION 1115069D01.x1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA
            sequence.
ACCESSION   CF561913
VERSION   CF561913.1   GI:34987996
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
ORGANISM   Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 422)
AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Merchant,S., Quinn,J. and Shrager,J.
TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1115
JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 422.
FEATURES   Location/Qualifiers
     source   1..422
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant (1998) Methods in
                     Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
                     micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
                     Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,
                     Merchant S.LaFontaine et al. (2002) Eukaryotic Cell,
                     1:736-757) in a shaking illuminated incubator (same
                     conditions as above). CC-1690 cells were grown to mid-log
                     phase in TAP medium in a shaking illuminated incubator to
                     a density of 8x10e6 cells/ml. The flask was transferred to
                     a shaking platform (200 rpm) at room temperature (23C) 12
                     micromole/m2/sec illumination and bubbled in a stoppered

```
REFERENCE
AUTHORS     Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Merchant,S., Quinn,J. and Shrager,J.
TITLE       Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1115
JOURNAL     Unpublished (2003)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 438.
FEATURES             Location/Qualifiers
     source          1..438
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant(1998) Methods in
                     Enzymology, 2997:263-279)in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
                     micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
                     Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,
                     Merchant S.LaFontaine et al. (2002) Eukaryotic Cell,
                     1:736-757) in a shaking illuminated incubator (same
                     conditions as above). CC-1690 cells were grown to mid-log
                     phase in TAP medium in a shaking illuminated incubator to
                     a density of 8x1oe6 cells/ml. The flask was transferred to
                     a shaking platform (200 rpm) at room temperature (23C) 12
                     micromole/m2/sec illumination and bubbled in a stoppered
                     flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour
                     prior to harvesting for RNA isolation (as per Quinn JM,
                     Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC
                     275:6080-6089); CC-1690 cells grown to mid-log phase
                     (3x1oe6 cells/ml) in TAP medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x1oe6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to
                     mid-log phase in HS medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x1oe6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified
                     from each sample, pooled and cDNA synthesized (see Shrager
                     et al, Plant Physiol. 131, 401-408 for details). The cDNA
                     was directionally cloned into lambda Zap II (Stratagene)
                     in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK-
                     plasmids were excised from the lambda ZAP clones by
                     superinfection with ExAssist (Stratagene) phage. The
                     library was normalized using method 4 described in Bonaldo
                     et al., (1996) Genome Research 6: 791-806."
```

```
                     flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour
                     prior to harvesting for RNA isolation (as per Quinn JM,
                     Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC
                     275:6080-6089); CC-1690 cells grown to mid-log phase
                     (3x1oe6 cells/ml) in TAP medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x1oe6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to
                     mid-log phase in HS medium in a shaking (150 rpm)
                     illuminated (70 mole photon/m2/sec) incubator at 27 C.
                     Cells were diluted to 1x1oe6 cells/ml, transferred to high
                     light (11000 mol photon/m2/s) with shaking (150 rpm) and
                     sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified
                     from each sample, pooled and cDNA synthesized (see Shrager
                     et al, Plant Physiol. 131, 401-408 for details). The cDNA
                     was directionally cloned into lambda Zap II (Stratagene)
                     in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK-
                     plasmids were excised from the lambda ZAP clones by
                     superinfection with ExAssist (Stratagene) phage. The
                     library was normalized using method 4 described in Bonaldo
                     et al., (1996) Genome Research 6: 791-806."
ORIGIN

Query Match             26.5%;  Score 274;  DB 7;  Length 422;
Best Local Similarity   79.3%;  Pred. No. 5.8e-69;
Matches   344;  Conservative   0;  Mismatches   75;  Indels  15;  Gaps  1;

Qy    128 CGACCCGTAACGATCAGCTTGGTGCTGGTCGGTTCGGTGGTTACCAGGTTAACCCGTACC 187
          |||||| | |||||||||| | |||||||||| |||||| ||||||||||||| ||||
Db      4 CGACCCATGAAAACCAACTGGGCGCTGGTGCTTTTGGTGGTTACCAGGTTACCCGTATG 63

Qy    188 TCGGTTTCGAAATGGTTATGACTGGCTGGCCGTATGGCGATATAAAGGCAGCGTTGACA 247
          | || |||||||||||| |||||| || ||||| ||  | |||||||||||||||| |
Db     64 TTGGCTTTGAAATGGTTACGACTGGTTAGTCGTATGTCGTATGCCGTACAAAGGCAGCGTTGAAA 123

Qy    248 ACGGYGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGGCTAAACTGGTTACCGGATCACTG 307
          |||| ||  | ||||||||||||||||| |||||||||| ||||||||||| |||||||
Db    124 ACGGTGCATACAAAGCTCAGGGCGTTCAACTGACCGGCGTTCAACTGGTTACCCAATCACTG 183

Qy    308 ACGATCTGGACATCTACACCCGTCGGGCGGCCATGGTTTGGCGCGGCTGACTCCAAAGGCA 367
          |||| |||||||||  |||| ||||| | ||||||| ||||| | |  | ||| |  |||
Db    184 ACGACCTGGACATCTCACACTCGTCGGGGTGGCATGGTATGGCGTGCGACACTAAATCCA 243

Qy    368 ACTACGCTTCTACCGGCGTTTCCCGTAGGCAGGCAACGACGACACTGGCGTTCCCAGTATTTG 427
          |||                    |||| ||| |||||||| ||| |||||| | ||   |||
Db    244 A----------CGTTTATGGTAAAAACCACGACACCGGCGGTCCGGTCTTCG 288

Qy    428 CTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGG 487
          ||||| | || ||||  || |||||||| || | | |||   |||||||||||||||| |||
Db    289 CTGGCGTGTTGAGTACGCGGATCACTCCTGAAATCGCTCGAAATCGCTGGAATACCAGTGGA 348

Qy    488 TTAACAACATCGGGCACGCGGGACACTGTGGGTACCCGTCCTGATAACGGCATGCTGAGCC 547
          |||| |||||||| | ||| || |||| || || ||||| ||| |||||||||| |||| |
Db    349 CCACAACAACTCGGTGACGCACGCACACCATCGGCACTCCTTCCGGACAACGGCCATGCTGAGTC 408

Qy    548 TGGGCGTTCCTAC 561
          ||||| |||||||
Db    409 TGGGTGTTCCTAC 422

RESULT 7
CF563205
LOCUS       CF563205                 438 bp    mRNA    linear   EST 23-SEP-2003
DEFINITION  1115082A08.x1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 3', mRNA
            sequence.
ACCESSION   CF563205
VERSION     CF563205.1  GI:34989288
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
ORGANISM    Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
```

```
Qy   753  TTCCGCTGTTGTTCTGGGTCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCT   812
          ||||| ||||||||||||| |||||||||||||||||||| || ||||||||||||||| |
Db   730  TTCCCGTAGTTGTTCTGGGTTTACACCGACCGCATCGGTTCTGACGCTTACAACCAGCAGGT   671

Qy   813  GTCTGAGAAACGTGCTCAGCTGCCGTTGTGTTGCTGGTGCTAAAGGCATCCCGCTGG   872
          ||| ||| | |||||||||||||||||||||||||      || ||||| ||||  |
Db   670  GTCCGAGCGCCGCCGTGCTCAGCTGCCGTTGTGTTGATTCCAAAGGTATCCCGGCAGA   611

Qy   873  CAAAATCTCCGCTGCGGCATGGGTGAATCAACCCGGTTACTGGCAACACCTGTGACAA   932
          ||| |||||||| ||| ||||| ||||||||||||||||||||||||||||||||||
Db   610  CAAGATCTCCGCACGTGGTATGGGCGAATCAACCCGGTTACTGGCAACACCTGTGACAA   551

Qy   933  CGTGAAAGTCGCGCTGCCCTGATCGATTGCTGGTCCCGATCGTCGTGTAGGATCGA   992
          ||||||| | |||||| ||| ||||||| |||||||||||||| |||| ||||||||
Db   550  CGTGAAACAGCGTGCTGCACTGATCGACTGCTGGTCCCGATCGTCCGGATCGAGATCGA   491

Qy   993  AGTTAAAGGCTACAAAGAGTTGTAACTCAGCCGGCGGGTTAA   1035
          |||||||||  ||||||| |||||||||||||| | |||||||
Db   490  AGTTAAAGGTATCAAAGACGTTGTAACTCAGCCGCCAGGCTTAA   448
```

RESULT 9
CF563206/c
LOCUS   CF563206   323 bp   mRNA   linear   EST 23-SEP-2003
DEFINITION 1115082A08.y1 C. reinhardtii CC-1690 (mt+), Stress III
            (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA 5', mRNA
            sequence.
ACCESSION   CF563206
VERSION   CF563206.1 GI:34989289
KEYWORDS   EST.
SOURCE   Chlamydomonas reinhardtii
  ORGANISM Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1 (bases 1 to 323)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Merchant,S., Quinn,J. and Shrager,J.
  TITLE   Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1115
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu
            High quality sequence stop: 323.
FEATURES   Location/Qualifiers
     source   1..323
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690 (mt+), Stress III
                     (normalized), Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; Stress condition III library, constructed by
                     Jeanette Quinn and Chiung-Wen Chang, combines cDNAs from
                     CC-1690 cells grown to mid-log phase in copper-free TAP
                     medium (see Quinn and Merchant(1998) Methods in
                     Enzymology, 2997:263-279) in a shaking(250 rpm) illuminated
                     (about 100 micromole/m2/sec) incubator at 22 C (see Quinn
                     and Merchant(1998) Methods in Enzymology, 2997:263-279);
                     CC-1690 cells grown to mid-log phase in low Fe (1
                     micromolar Fe) TAP medium (see La Fontaine S, Quinn JM,
                     Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J,
                     Merchant S.LaFontaine et al. (2002) Eukaryotic Cell,
                     1:736-757) in a shaking illuminated incubator (same
                     conditions as above). CC-1690 cells were grown to mid-log
                     phase in TAP medium in a shaking illuminated incubator to
```

```
Qy   190  GGTTTCGAAATGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAAC   249
          || || |||||||| ||||| ||||| ||  ||| ||| ||| |||||||||||||| ||
Db    87  GGCTTTGAAATGGGTTACGACTGGTTAGGTCGTATGCCCTACACAAAGGCAGCGTTGAAAAC   146

Qy   250  GGTGCTTTCAAAGTCAGGCGGTTCAGCTGACCGTTCAGCTGACCGTTACCGGATCACTGAC   309
          |||||  | ||||| ||||| |||||| |||||||||||| |||||||||||| |||||||
Db   147  GGTGCATACAAAGCTCAGGGCGGTTCAACTGACCGTTCAACTGACCGTTACCAATCACTGAC   206

Qy   310  GATCTGGACATCTACACCGTCTCGGGCGCCATGGTTTGGCGCGCCTGACTCCAAAGGCAAC   369
          || |||||||||||||||||||| || ||||||  | ||||||| |||||| |||  |||
Db   207  GACCTGGACATCTACACCGTCTCTGGTGCCATGGCATGCGTGCTGCAGACAACTAAATCCAA-   265

Qy   370  TACGCTTCTACCGGCGGTTTCCCGTAGCGAACGACACACTGGCGTTTCCCAGTATTTGCT   429
          |||| ||             |||||      ||||||||| |||||| |||| ||| |||
Db   266  --------CGTTTATGTGAAAAACCCACGACACACCGGCGTTCTCCGGTCTTCGCT   311

Qy   430  GGCGGCCGTAGAGTGGGCGTGTTACTCGTGGACATCGCTCGCTCGGAATACCAGTGGGTT   489
          ||||| ||||||| ||||| | ||||  |||||||||||| |||||||||||||||| ||
Db   312  GGCGGTAGTGAGTACGGCGATCATCGCTCTCGAAATCGCTCGGAATACCAGTGGACC   371

Qy   490  AACAACATCGGCGACGCGGGGCACTGTGGGTACCCGTCCTGATAACGGCATGCTGAGCCTG   549
          ||||||||||| |||||| |  |  || |   ||||||| ||| |||||||||||||||||
Db   372  AACAACATCGGTGACGCGGACGCACACCATCGGCACTCGTCCGGACACGGCATGCTGAGCCTG   431

Qy   550  GGCGTTT   556
          || ||||
Db   432  GGTGTTT   438
```

RESULT 8
CL657503/c
LOCUS   CL657503   730 bp   DNA   linear   GSS 09-JUL-2004
DEFINITION PRI0129d_A09 - PRI0129d.B21 (730) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL657503
VERSION   CL657503.1 GI:50139169
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 730)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE   AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT   Contact: Sommer RJ
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES   Location/Qualifiers
     source   1..730
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

Query Match   21.2%;   Score 219;   DB 9;   Length 730;
Best Local Similarity   85.9%;   Pred. No. 9.4e-53;
Matches 243; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

a density of 8x10e6 cells/ml. The flask was transferred to a shaking platform (200 rpm) at room temperature (23C) 12 micromole/m2/sec illumination and bubbled in a stoppered flask with 98% nitrogen, 2% CO2 gas mixture for 1 hour prior to harvesting for RNA isolation (as per Quinn JM, Barraco P, Eriksson M, Merchant S. Quinn et al. (2000) JBC 275:6080-6089).; CC-1690 cells grown to mid-log phase (3x10e6 cells/ml) in TAP medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs); CC-1690 cells grown to mid-log phase in HS medium in a shaking (150 rpm) illuminated (70 mole photon/m2/sec) incubator at 27 C. Cells were diluted to 1x10e6 cells/ml, transferred to high light (11000 mol photon/m2/s) with shaking (150 rpm) and sampled at (0.5, 1,2,4,6, 12 hrs). PolyA mRNA was purified from each sample, pooled and cDNA synthesized (see Shrager et al, Plant Physiol. 131, 401-408 for details). The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK-plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

ORIGIN

Query Match    17.0%;    Score 176;    DB 7;    Length 323;
Best Local Similarity    84.9%;    Pred. No. 3.1e-40;
Matches 197;    Conservative 0;    Mismatches 35;    Indels 0;    Gaps 0;

```
Qy  804 CCAGCAGCTGCTGCTGAGAAACGTGCTCCAGTCCGTTGTGTTGACTACCTGGTTGCTAAAGGCAT  863
            |||| |||| |||  ||  ||| ||  ||||| |||| || ||  ||||| ||||  ||||| |
Db  323 CCAGGGTCTGCCGAGCGCGCCCGTGCTGTGTTGATTACCTGATCTCCAAAGGTAT  264

Qy  864 CCCGGCTGGCAAAATCTCCGCTGCGCGCAGCATGGTGAATCCAACCCGGTTACTGGCAACAC  923
            ||||||||   |  ||| ||| |||||||  |||| ||| |||||  || |||||| |||||||
Db  263 CCCGGCAGACAAGATCTTCCGCACGTGCTATGGGCGAATCCAACCCACCCGGTTACTGGCAACAC  204

Qy  924 CTGTGACAACGTGAAAGCTGCGCGTCGCCTGCCCTGATCGATTGCCTGGCTCCGGATCGTCGTGT  983
            |||||||||  || ||| ||  |  ||| |||||  |||||||  || |||||||  ||||||||
Db  203 CTGTGACAACGTGAAACAGCAGCGGTGCACTGACTGCTGCGTTCTCCGGATCGTCGCGT  144

Qy  984 AGAGGATCGAAGTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCCGGGTTAA  1035
            ||||||||||||||||| |  ||| ||||||||||||||||| ||||| |||
Db  143 AGAGGATCGAAGTTAAAGGTATCAAAGACGTGTAACTCAGCCGCCGCAGGCTTAA  92
```

RESULT 10
BE423296
LOCUS   BE423296   603 bp   mRNA   linear   EST 24-JUL-2000
DEFINITION   WHE0063_E04_I07zS Wheat endosperm cDNA library Triticum aestivum cDNA clone WHE0063_E04_I07, mRNA sequence.
ACCESSION   BE423296
VERSION   BE423296.1   GI:9421139
KEYWORDS   EST.
SOURCE   Triticum aestivum
ORGANISM   Triticum aestivum (bread wheat)
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 603)
AUTHORS   Altenbach,S., Anderson,O.D., Chao,S., Galili,G., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
TITLE   The structure and function of the expressed portion of the wheat genomes - Endosperm cDNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Olin Anderson
             US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
             800 Buchanan Street, Albany, CA 94710, USA
             Tel: 5105595773

Fax: 5105595818
Email: oandersrn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES   Location/Qualifiers
source   1..603
            /organism="Triticum aestivum"
            /mol_type="mRNA"
            /cultivar="Cheyenne"
            /db_xref="taxon:4565"
            /clone="WHE0063_E04_I07"
            /tissue_type="Endosperm"
            /dev_stage="5 to 30 days post anthesis seed"
            /lab_host="E. coli SOLR"
            /clone_lib="Wheat endosperm cDNA library"
            /note="Vector: Lambda ZAP II, excised phagemid; Site_1: EcoRI; Seeds collected, endosperm isolated, and RNA prepared by Susan Altenbach. Library constructed by Stratagene, Inc. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab."

ORIGIN

Query Match   15.4%;   Score 159;   DB 2;   Length 603;
Best Local Similarity   77.9%;   Pred. No. 3.6e-35;
Matches 211;   Conservative 0;   Mismatches 45;   Indels 15;   Gaps 1;

```
Qy  200 TGGGTTATGACTGGCTGGCCGTATGCGCATATAAAGGCAGCCTTGACAACGGTGCTTTCA  259
            |||||| ||| ||| ||| |||| |||   |||||||||||||||| |||||||||| |
Db  149 TGGGTTACGACTGGTTAGGTCGTATCGTACCGTACAAAGGCAGCGTTGAAAACGGTGCATACA  208

Qy  260 AAGCTTCAGGGCGTTCAGCTGACCGCTCAAACTGGGTTACCGGTCACTGACGATCTGACA  319
            |||||||||||||| |||  |||||||||||| ||||| |||| |||| ||| |||||||
Db  209 AAGCTTCAGGGCGTTCAACTGACCGCTCAAACTGGGTTACCCATCACTGACGACCTGACA  268

Qy  320 TCTACACCCGTCTGGGCGGCATGGTTGGCGCGCTACTCCAAAGGCAACTACGCTTCTA  379
            ||||||| ||| || || ||||| ||| || ||| ||  ||||| ||| ||
Db  269 TCTACACTCGTCTGGGTGGCATGGTATGGTGCGAGCACTAAATCCAA----------  317

Qy  380 CCGGCGGTTCCCGTAGCGAACACGACACTGGCGTTTCCCCAGTATTTGCTGGCGGCGTAG  439
            |||||||||    |  ||| |  ||| |||  |||||  ||  ||| ||||| ||||||  |
Db  318 ----CGTTTATTGGTAAAAACCACGACACCGACGCGTTTCTCCGGTCTCGCTGGCGTGTTG  373

Qy  440 AGTGGCTGTTACTCGTGACATCGCTACCCG  470
            || || |||| |||| ||||||| ||||||
Db  374 AGTACGCGGATCACTCCTGAAATCGCTCGCCG  404
```

RESULT 11
AF075794/c
LOCUS   AF075794   655 bp   DNA   linear   GSS 29-AUG-2000
DEFINITION   AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 107-T3, genomic survey sequence.
ACCESSION   AF075794
VERSION   AF075794.1   GI:3320664
KEYWORDS   GSS.
SOURCE   Salmonella typhimurium
ORGANISM   Salmonella typhimurium
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.
REFERENCE   1 (bases 1 to 655)
AUTHORS   Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
TITLE   Sample sequencing of a Salmonella typhimurium LT2 lambda library: comparison to the Escherichia coli K12 genome
JOURNAL   FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE   99243757
PUBMED   10227170
COMMENT   Contact: McClelland M
             Molecular Biology
             Sidney Kimmel Cancer Center
             3099 Science Park Road, San Diego, CA 92121, USA
             Email: mcclelland@lifsci.sdsu.edu
             Class: shotgun.

```
FEATURES             Location/Qualifiers
     source          1..655
                     /organism="Salmonella typhimurium"
                     /mol_type="genomic DNA"
                     /strain="LT2"
                     /db_xref="taxon:602"
                     /clone="107-T3"
                     /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
                     /note="Vector: Lambda DASH II; sequenced using Li-Cor
                     sequencer"

ORIGIN

Query Match             13.0%;  Score 134.2;  DB 8;  Length 655;
Best Local Similarity   72.7%;  Pred. No. 7.1e-28;
Matches   194;  Conservative   0;  Mismatches  58;  Indels  15;  Gaps   1;

Qy    15 CGTACTGAATGCGGCTCCGAAAGATAACACCTGTATGCAGGTTGGTAAACTGGGTTGGTC 74
            ||| | |||| ||||||||||||||||||||| |||||||  ||||||| ||| |||||
Db   252 CGTAGCGCAGGCCCCTCCGAAAGATAACACCTGTGGTCTAAACTGGGCTGGTC 193

Qy    75 CCAGTATCACGACCACCGGTTTCTACGGTAACGGTTTCCAGACAACAACGGTCCGACCCG 134
            |||||||       ||||   ||||  ||||||  ||||||||
Db   192 TCAGTACCATGACAC------------CGGCTTCATTCACAATGATGGCCCGACTCA 148

Qy   135 TAACGATCAGCTTGGTGCGTGCGTTCGGTGGTTACCAGTTAACCCGTACCTCGGTTT 194
            | | ||| || ||||||||   | ||||||| | ||||||  || ||| |||||
Db   147 TGAAAACCAACTGGGCGCCAGGTGCTTTGGTGGTGGTTACCAGTTAACCCGTATGTTGGCTT 88

Qy   195 CGAAATGGGTTATGACTGGCTGGGCCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGC 254
            ||||||| || ||||||||| ||| ||||||| |||||||||||| ||| ||||| ||
Db    87 TGAAATGGGCTACGACTGTTAGCCGTATGCCGTATGCCGTACAAAGGCGACTACACTCAATGGGCGC 28

Qy   255 TTTCAAAGCTCAGGGCGGTTCAGCTGAC 281
            |||  |||||||||||  |||||||||||
Db    27 TTATAAAGCTCAGCGCGGTTCAGTTGAC 1

RESULT 12
CF915072
LOCUS       CF915072                 601 bp    mRNA    linear     EST 05-NOV-2003
DEFINITION  B0973F11-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
            musculus cDNA clone NIA:B0973F11 IMAGE:30479302 5', mRNA sequence.
ACCESSION   CF915072
VERSION     CF915072.1  GI:38186274
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 601)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
  MEDLINE   21429098
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: B0973 row: F column: 11
            Seq primer: M13 Reverse
            High quality sequence stop: 601
            POLYA=No.

FEATURES             Location/Qualifiers
     source          1..601
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:B0973F11-5"
                     /db_xref="taxon:10090"
                     /clone="NIA:B0973F11 IMAGE:30479302"
                     /dev_stage="Unfertilized Egg"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
                     1)"
                     /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
                     Site_2: NotI; Mouse cDNA project by the Laboratory of
                     Genetics, National Institute on Aging (NIA), Intramural
                     Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
                     This is a long-transcript enriched cDNA library (Ref.
                     Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
                     RNAs were extracted from a pool of 1488 unfertilized eggs.
                     Double-stranded cDNAs were synthesized with an Oligo(dT)
                     primer [Invitrogen]:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCTTTTTTTTTTTTTTTTTTT-3',
                     treated with T4 DNA polymerase, and purified by
                     ethanol-precipitation. The cDNAs were ligated to
                     Lone-linker LL-Sal4, purified by phenol/chloroform, and
                     separated from free linkers by Centricon 100. Then, the
                     cDNAs were amplified by long-range high fidelity PCR using
                     Ex Taq polymerase (Takara) with a primer Sal4-S. The
                     products were purified by phenol/chloroform and Centricon
                     100. The cDNAs were digested with SalI and NotI enzymes
                     and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
                     vector. The DH10B E. coli host was transformed with the
                     ligation mixture by the standard chemical method. The
                     average insert size is about 2.5 kb. The library was
                     constructed by Yulan Piao."

ORIGIN

Query Match             5.9%;   Score 61.4;   DB 7;   Length 601;
Best Local Similarity   59.4%;  Pred. No. 1.6e-06;
Matches   104;  Conservative   0;  Mismatches  71;  Indels   0;  Gaps   0;

Qy   755 CCGCTGTGTTGTTCTGGGCTACACCGACCCGATCGTTCCGAAGCTTACACCAGCAGCTGT 814
            || ||||| ||||||| ||||||| |||||| |||| ||| ||||||||||||
Db   339 CCACCGTTGTTGAAGGTCACACTGACTCCGTCGGTCCTGACGCTTACACCAGCAGAGCTGT 398

Qy   815 CTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCA 874
            ||||| | ||| |||||||| || ||||| ||  ||||  |||   |||||||
Db   399 CCCAGCGTCGTGCTGACGCTGTTAAGCAAGTCCTGGTCAAAGACGGTATTGCTCCTAACC 458

Qy   875 AAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGTGA 929
            | |||||| ||||||| ||||||| ||| |||||||||||   || || ||||||
Db   459 GTGTAAGCTCGGTTGGTTATGGCAATATCCCGCCCAGTGCTGACAACGCAACTGA 513

RESULT 13
CF915603
LOCUS       CF915603                 601 bp    mRNA    linear     EST 05-NOV-2003
DEFINITION  B0982E05-5 NIA Mouse Unfertilized Egg cDNA Library (Long 1) Mus
            musculus cDNA clone NIA:B0982E05 IMAGE:30480148 5', mRNA sequence.
ACCESSION   CF915603
VERSION     CF915603.1  GI:38186805
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 601)
  AUTHORS   Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
  TITLE     Construction of long-transcript enriched cDNA libraries from
            submicrogram amounts of total RNAs by a universal PCR amplification
            method
  JOURNAL   Genome Res. 11 (9), 1553-1558 (2001)
  MEDLINE   21429098
   PUBMED   11544199
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: B0982 row: E column: 05
```

```
JOURNAL      Unpublished (2003)
COMMENT      Contact: Blaxter ML
             Institute of Cell, Animal and Population Biology
             University of Edinburgh
             Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
             3JT, UK.
             Tel: +44 131 650 6760
             Fax: +44 131 670 5450
             Email: mark.blaxter@ed.ac.uk
             The library was constructed from mixed stages of the tardigrade
             Hypsibius dujardini maintained in lab culture. The original source
             culture was obtained from Sciento (strain Z151). The library was
             constructed by Jennifer Daub, University of Edinburgh. Sequencing
             was performed by Frances Thomas, University of Edinburgh.
             PCR PRimers
             FORWARD: M13R
             BACKWARD: M13L
             Plate: 67  row: B  column: 10
             Seq primer: T7 (TAATACGACTCACTATAGGG)
             High quality sequence start: 6
             High quality sequence stop: 533.

FEATURES     Location/Qualifiers
    source   1..736
             /organism="Hypsibius dujardini"
             /mol_type="mRNA"
             /db_xref="taxon:232323"
             /clone="Hd_mx17_67B10"
             /sex="parthenogenic"
             /tissue_type="whole body"
             /dev_stage="mixed (adult and juvenile)"
             /clone_lib="Hypsibius dujardini mixed stage (fraction 7)"
             /note="Vector: pSPORT1; Site_1: Sal I (5'end); Site_2: Not
             I (3'end); Tardigrades (also know as water bears) are
             small free living animals inhabiting marine, fresh water
             and water film habitats. Hypsibius dujardini is a
             freshwater species maintained in lab culture. The library
             was prepared from adults and juveniles and was constructed
             by Jennifer Daub, University of Edinburgh."

ORIGIN

Query Match       5.9%;   Score 61.4;  DB 7;  Length 736;
Best Local Similarity  59.4%;  Pred. No. 1.7e-06;
Matches  104;  Conservative   0;  Mismatches  71;  Indels   0;  Gaps   0;

Qy   755 CCGCTGTTGTTCTGGGCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCTGT 814
             ||||||||||||||||||||||| ||||| |||||||||||||||||||||||||||||
Db   357 CCACCCGTTGTTGAAGGTCACACTGACTCCGTCCGGTCCTGGTTGCTAAAGGCATCCCGGCTGCA 416

Qy   815 CTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCA 874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   417 CCCAGCGTCGTGCTGACGCTGTTAAGCAAGTCCTGGTCAAAGACGGTATTGCTCCTAACC 476

Qy   875 AAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGTGA 929
             ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   477 GTGTAAGCTCGTCGGTTGGTTATGCAAATCCCGCCCAGTGTGCTGCTGACAACGCAACTGA 531

RESULT 15
CK326686
LOCUS      CK326686              763 bp   mRNA   linear   EST 17-DEC-2003
DEFINITION  Hd_mx24_02F09 T7 Hypsibius dujardini mixed stage 5' capped (fraction
            4) Hypsibius dujardini cDNA clone Hd_mx24_02F09 5' similar to
            Q52465 (Q52465) Protein F precursor (Fragment), mRNA sequence.
ACCESSION   CK326686
VERSION     CK326686.1  GI:40006296
KEYWORDS    EST.
SOURCE      Hypsibius dujardini
  ORGANISM  Hypsibius dujardini
            Eukaryota; Metazoa; Tardigrada; Eutardigrada; Apochela;
            Hypsibiidae; Hypsibius.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   Daub,J., Thomas,F., Aboobaker,A. and Blaxter,M.L.
  TITLE     A survey of genes expressed in the tardigrade Hypsibius dujardini
```

```
FEATURES     Seq primer: M13 Reverse
             High quality sequence stop: 601
             POLYA=No.

    source   1..601
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6J"
             /db_xref="niaEST:B0982E05-5"
             /clone="NIA:B0982E05 IMAGE:30480148"
             /dev_stage="Unfertilized Egg"
             /lab_host="DH10B"
             /clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long
             1)"
             /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
             Site_2: NotI; Mouse cDNA project by the Laboratory of
             Genetics, National Institute on Aging (NIA), Intramural
             Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
             This is a long-transcript enriched cDNA library (Ref.
             Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
             RNAs were extracted from a pool of 1488 unfertilized eggs.
             Double-stranded cDNAs were synthesized with an Oligo(dT)
             primer [Invitrogen]:
             5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTTTT-3'],
             treated with T4 DNA polymerase, and purified by
             ethanol-precipitation. The cDNAs were ligated to
             Lone-linker LL-Sal4, purified by phenol/chloroform, and
             separated from free linkers by Centricon 100. Then, the
             cDNAs were amplified by long-range high fidelity PCR using
             Ex Taq polymerase (Takara) with a primer Sal4-S. The
             products were purified by phenol/chloroform and Centricon
             100. The cDNAs were digested with SalI and NotI enzymes
             and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
             vector. The DH10B E. coli host was transformed with the
             ligation mixture by the standard chemical method. The
             average insert size is about 2.5 kb. The library was
             constructed by Yulan Piao."

ORIGIN

Query Match       5.9%;   Score 61.4;  DB 7;  Length 601;
Best Local Similarity  59.4%;  Pred. No. 1.6e-06;
Matches  104;  Conservative   0;  Mismatches  71;  Indels   0;  Gaps   0;

Qy   755 CCGCTGTTGTTCTGGGCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCTGT 814
             ||||||||||||||||||||||| ||||| |||||||||||||||||||||||||||||
Db   339 CCACCCGTTGTTGAAGGTCACACTGACTCCGTCCGGTCCTGACGCTTACAACCAGAAGCTGT 398

Qy   815 CTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCA 874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   399 CCCAGCGTCGTGCTGACGCTGTTAAGCAAGTCCTGGTCAAAGACGGTATTGCTCCTAACC 458

Qy   875 AAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGTGA 929
             ||| ||||||||| |||||||||||||||||||||||||||||||||||||||
Db   459 GTGTAAGCTCGTCGGTTGGTTATGCAAATCCCGCCCAGTGTGCTGCTGACAACGCAACTGA 513

RESULT 14
CF544559
LOCUS      CF544559              736 bp   mRNA   linear   EST 01-OCT-2003
DEFINITION  Hd_mx17_67B10 T7 Hypsibius dujardini mixed stage (fraction 7)
            Hypsibius dujardini cDNA clone Hd_mx17_67B10 5' similar to Q52465
            (Q52465) Protein F precursor (Fragment), mRNA sequence.
ACCESSION   CF544559
VERSION     CF544559.1  GI:37213423
KEYWORDS    EST.
SOURCE      Hypsibius dujardini
  ORGANISM  Hypsibius dujardini
            Eukaryota; Metazoa; Tardigrada; Eutardigrada; Apochela;
            Hypsibiidae; Hypsibius.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Daub,J., Thomas,F., Aboobaker,A. and Blaxter,M.L.
  TITLE     A survey of genes expressed in the tardigrade Hypsibius dujardini
```

JOURNAL   Unpublished (2003)
COMMENT   Contact: Blaxter ML
             Institute of Cell, Animal and Population Biology
             University of Edinburgh
             Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
             3JT, UK.
             Tel: +44 131 650 6760
             Fax: +44 131 670 5450
             Email: mark.blaxter@ed.ac.uk
             PCR Primers
             FORWARD: M13R
             BACKWARD: M13L
             Plate: 02 row: F column: 09
             Seq primer: T7 (TAATACGACTCACTATAGGG)
             High quality sequence stop: 542.
FEATURES   Location/Qualifiers
             source   1..763
                             /organism="Hypsibius dujardini"
                             /mol_type="mRNA"
                             /db_xref="taxon:232323"
                             /clone="Hd_mx24_02F09"
                             /sex="parthenogenic"
                             /tissue_type="whole body"
                             /dev_stage="mixed (adult and juvenile)"
                             /clone_lib="Hypsibius dujardini mixed stage 5'capped
                             (fraction 4)"
                             /note="Vector: pSPORT1; Site_1: Sal I (5'end); Site_2: Not
                             I (3'end); Tardigrades (also know as water bears) are
                             small free living animals inhabiting marine, fresh water
                             and water film habitats. Hypsibius dujardini is a
                             freshwater species maintained in lab culture (source
                             strain Sciento Z151). The library was prepared from adults
                             and juveniles and was constructed in pSPORT1 vector
                             (Superscript, invitrogen) using GeneRacer (Invitrogen)
                             purified 5'capped mRNA by Jennifer Daub, University of
                             Edinburgh. Sequencing was performed by Frances Thomas,
                             University of Edinburgh."

ORIGIN

Query Match   5.9%;   Score 61.4;   DB 7;   Length 763;
Best Local Similarity 59.4%;   Pred. No. 1.7e-06;
Matches 104;   Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 755 CCGCTGTGTTCTGGGCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCTGT 814
            |||||| ||||||||| | | ||||| |||| |||| ||||||| |||||||||||
Db 356 CCACCGTTGTTGAAGGTCACACTGACTCCGTCCGTCGTGACCGCTTACAACCAGAAGCTGT 415

Qy 815 CTGAGAAACGTGCTCAGTCCGTTGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCA 874
            || |||| |||||| || |||| | | || || ||| |||| |||
Db 416 CCCAGCGTCGTGCTGACGCTGTTAAGCAAGTCCTGGTCAAAGACGGTATTGCTCCTAACC 475

Qy 875 AAATCTCCGCTCCGCCGGCCATGGGTGAATCAACCCGGTTACTGGCAACACCTGTGA 929
            | | | || ||| || ||| | | | |||||||| |||| ||| |||
Db 476 GTGTAAGCTCGGTTGGTTATGGCAAATCCCGCCCAGTTGCTGACAACCCAACTGA 530

Search completed: January 19, 2005, 16:36:47
Job time : 3915 secs

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005 Compugen Ltd.
```

Run on: January 19, 2005, 15:31:38 ; Search time 669 Seconds
                                          (without alignments)
                                          8889.379 Million cell updates/sec

Title: US-09-913-772-1
Perfect score: 1035
Sequence: 1 atgaaagcaatttcgtact........taactcagccggcgggttaa 1035

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 287294193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA:*
        1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
        2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
        3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
        4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
        5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
        6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
        7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
        8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
        9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
       10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
       11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
       12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
       13: /cgn2_6/ptodata/2/pubpna/US10A_NEW_PUB.seq:*
       14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
       15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
       16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
       17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
       18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
       19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
       20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
       21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1035	100.0	1035	14	US-10-169-953-1	Sequence 1, Appli
2	1008	97.4	1008	14	US-10-091-257-13	Sequence 13, Appl
3	724.4	70.0	730	9	US-09-452-599-11	Sequence 11, Appl
4	724.4	70.0	730	15	US-10-121-120-11	Sequence 11, Appl
5	696.2	67.3	1422	18	US-10-425-821-69	Sequence 69, Appl
6	403	38.9	411	16	US-10-432-056-1	Sequence 1, Appli
7	303.4	29.3	3630	17	US-10-416-708A-23	Sequence 23, Appl
8	300.6	29.0	3660	17	US-10-416-708A-26	Sequence 26, Appl
9	298.6	28.9	3660	17	US-10-416-708A-9	Sequence 9, Appli
10	298.6	28.9	4543	17	US-10-416-708A-62	Sequence 62, Appl
11	219	21.2	3817	15	US-10-033-399B-24	Sequence 24, Appl
12	161	15.6	1095	11	US-09-809-665A-150	Sequence 150, App
13	161	15.6	1095	18	US-10-854-299-150	Sequence 150, App
14	159.4	15.4	1922	18	US-10-777-824-7	Sequence 7, Appli
15	135	13.0	1035	15	US-10-336-840-20	Sequence 20, Appl
16	133.4	12.9	1035	15	US-10-336-840-17	Sequence 17, Appl
17	133.4	12.9	1035	15	US-10-336-840-21	Sequence 21, Appl
18	133.4	12.9	1035	15	US-10-336-840-23	Sequence 23, Appl
19	130	12.6	1830121	14	US-10-329-960-1	Sequence 1, Appli
20	130	12.6	1830121	16	US-10-329-670-1	Sequence 1, Appli
21	130	12.6	1830121	18	US-10-158-865-1	Sequence 1, Appli
22	129.8	12.5	1319	18	US-10-770-824-9	Sequence 9, Appli
23	127.2	12.3	1059	17	US-10-467-421-20	Sequence 20, Appl
24	125.4	12.1	1035	15	US-10-336-840-15	Sequence 15, Appl
25	125.4	12.1	1035	15	US-10-336-840-16	Sequence 16, Appl
26	125.4	12.1	1035	15	US-10-336-840-18	Sequence 18, Appl
27	125.4	12.1	1035	15	US-10-336-840-24	Sequence 24, Appl
28	125	12.1	1110	11	US-09-809-665A-152	Sequence 152, App
29	125	12.1	1110	18	US-10-854-299-152	Sequence 152, App
30	117.2	11.3	1026	15	US-10-336-840-19	Sequence 19, Appl
31	117.2	11.3	1026	15	US-10-336-840-22	Sequence 22, Appl
32	117.2	11.3	1026	15	US-10-336-840-25	Sequence 25, Appl
33 c	69.4	6.7	640681	9	US-09-790-988-1	Sequence 1, Appli
34	64.6	6.2	1349	17	US-10-467-534-100	Sequence 100, App
35	63	6.1	729	17	US-10-467-421-7	Sequence 7, Appli
36	61.4	5.9	729	17	US-10-467-421-9	Sequence 9, Appli
37	49.6	4.8	786	16	US-10-389-647-231	Sequence 231, App
38	44.8	4.3	687	16	US-10-282-122A-33267	Sequence 33267, A
39 c	42.4	4.1	2054	16	US-10-425-114-22314	Sequence 22314, A
40	41.8	4.0	1395	15	US-10-369-493-35123	Sequence 35123, A
41	41.8	4.0	1395	15	US-10-369-493-38184	Sequence 38184, A
42	41.8	4.0	1395	15	US-10-369-493-38576	Sequence 38576, A
43	40.6	3.9	1439	15	US-10-193-002-162	Sequence 162, App
44	40.6	3.9	1439	15	US-10-084-843-167	Sequence 167, App
45	40.2	3.9	678	15	US-10-156-761-3232	Sequence 3232, Ap

ALIGNMENTS

```
RESULT 1
US-10-169-953-1

; Sequence 1, Application US/10169953
; Publication No US20030044915A1
; GENERAL INFORMATION:
; APPLICANT: Thierry BAUSSANT
; APPLICANT: Pascale JEANNIN
; APPLICANT: Yves DELNESTE
; APPLICANT: FranOois LAWNY
; APPLICANT: Jean-Yves BONNEFOY
; TITLE OF INVENTION: METHOD FOR PREPARING A POLYPEPTIDE SOLUBLE IN AN AQUEOUS SOLVENT
; TITLE OF INVENTION: IN THE ABSENCE OF DETERGENT
; FILE REFERENCE: D18390
; CURRENT APPLICATION NUMBER: US/10/169,953
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: FR 00 00070
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
US-10-169-953-1
; SEQ ID NO 1
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1032)
; FEATURE:
;  OTHER INFORMATION: P40
US-10-169-953-1

Query Match             100.0%;  Score 1035;  DB 14;  Length 1035;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1035;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAAGCAATTTCGTACTACTACTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGTGGT 60
```

US-10-091-257-13
; Sequence 13, Application US/10091257
; Publication No. US20030064078A1
GENERAL INFORMATION:
‑ APPLICANT: Binz, Hans
‑ N'Guyen, Ngoc Thien
‑ Baussant, Thierry
‑ Trudel, Michel
TITLE OF INVENTION: PEPTIDE FRAGMENT OF RESPIRATORY
SYNCYTIAL VIRUS PROTEIN G, IMMUNOGENIC AGENT, PHARMACEUTICAL
COMPOSITION CONTAINING IT AND PREPARATION PROCESS
‑ NUMBER OF SEQUENCES: 75
‑ CORRESPONDENCE ADDRESS:
‑ ADDRESSEE: Gordon W. Hueschen
‑ STREET: 715 The "H" Bldg., 310 East Michigan
Avenue
‑ CITY: Kalamazoo
‑ STATE: MI
‑ COUNTRY: USA
‑ ZIP: 49007
‑ COMPUTER READABLE FORM:
‑ MEDIUM TYPE: Floppy disk
‑ COMPUTER: IBM PC compatible
‑ OPERATING SYSTEM: PC-DOS/MS-DOS
‑ SOFTWARE: PatentIn Release #1.0, Version #1.30
(EPO)
‑ CURRENT APPLICATION DATA:
‑ APPLICATION NUMBER: US/10/091,257
‑ FILING DATE: 05-Mar-2002
‑ CLASSIFICATION: <Unknown>
‑ PRIOR APPLICATION DATA:
‑ APPLICATION NUMBER: US/08/721,979A
‑ FILING DATE: October 4, 1996
‑ APPLICATION NUMBER: FR 94 04009
‑ FILING DATE: 06-APR-1994
‑ ATTORNEY/AGENT INFORMATION:
‑ NAME: Hueschen, Gordon W.
‑ REGISTRATION NUMBER: 16,157
‑ REFERENCE/DOCKET NUMBER: PF57PCTUS/dln
‑ TELECOMMUNICATION INFORMATION:
‑ TELEPHONE: 616-382-0030
‑ TELEFAX: 616-382-2030
INFORMATION FOR SEQ ID NO: 13:
‑ SEQUENCE CHARACTERISTICS:
‑ LENGTH: 1008 base pairs
‑ TYPE: nucleic acid
‑ STRANDEDNESS: single
‑ TOPOLOGY: linear
‑ MOLECULE TYPE: peptide
‑ FEATURE:
‑ NAME/KEY: CDS
‑ LOCATION: 1..1008
‑ OTHER INFORMATION: /note= "name : P40"
‑ SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-091-257-13

Query Match   97.4%;   Score 1008;   DB 14;   Length 1008;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1008;   Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   28 GCTCCGAAAGATAACACCTGGTATGCAGGTGGTTGGTTGGTTAAACTGGGTTCCCAGTATCACGAC 87
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCTCCGAAAGATAACACCTGGTATGCAGGTGGTTGGTTGGTTAAACTGGGTTCCCAGTATCACGAC 60

Qy   88 ACCGGTTTCTACGGTGCGTTCCAGAACAACAACAACGGTCCGACCCGTAACCCGTAACGATCAGCTT 147
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACCGGTTTCTACGGTGCGTTCCAGAACAACAACAACGGTCCGACCCGTAACCCGTAACGATCAGCTT 120

Qy 148 GGTGCTGGTGGCGTTCGGTGGTTACCCGGTTAACCGTAACCCGTACCCGTTTCGAAATGGGTTAT 207
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GGTGCTGGTGGCGTTCGGTGGTTACCCGGTTAACCGTAACCCGTACCCGTTTCGAAATGGGTTAT 180

Qy 208 GACTGGCTGGGCCGTATGGCCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG 267

RESULT 2

Db   1 ATGAAAGCAATTTCGTACTACCAGTATCGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 AAACTGGGTTGGTTGGTTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
Db   61 AAACTGGGTTGGTTGGTTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy 121 ACGGTCCGACCCGTAACGATCAGCTTGGTGGTCGTGGTCGTGGTTACCAGGTTAAC 180
Db 121 ACGGTCCGACCCGTAACGATCAGCTTGGTGGTCGTGGTCGTGGTTACCAGGTTAAC 180

Qy 181 CCGTACCCTGACTCC 300
Db 181 CCGTACCCTGACTCC 300

Qy 301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCGGCATGGTTTGGGCCGCGTTGACTCC 360
Db 301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCGGCATGGTTTGGGCCGCGTTGACTCC 360

Qy 361 AAAGGCAACTACGCTTCTACCGGCGTTTCCGTAGCGAACACGACACTGGCCGTTTCCCCA 420
Db 361 AAAGGCAACTACGCTTCTACCGGCGTTTCCGTAGCGAACACGACACTGGCCGTTTCCCCA 420

Qy 421 GTATTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATGCGTACCCGTTGTTGCTCCG 480
Db 421 GTATTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATGCGTACCCGTTGTTGCTCCG 480

Qy 481 CAGTGGGTTAACACACATCGGCGACCGGGCACTGTGGTACCCGTCCTGATAACGCATG 540
Db 481 CAGTGGGTTAACACACATCGGCGACCGGGCACTGTGGTACCCGTCCTGATAACGCATG 540

Qy 541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCTGACACCGGTTGTTGCTCCG 600
Db 541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCTGACACCGGTTGTTGCTCCG 600

Qy 601 GCTCCGCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGGTTCTG 660
Db 601 GCTCCGCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGGTTCTG 660

Qy 661 TTCAACTTCAACAAAGCTACACCCTGAAACCGGAAGCTCTCTGGATCAGCTGTAC 720
Db 661 TTCAACTTCAACAAAGCTACACCCTGAAACCGGAAGCTCTCTGGATCAGCTGTAC 720

Qy 721 ACTCAGCTGAGCAACATGACGGATCCGAAAGACGGTTCCGCTGTGTTGTCGGGCTACACCGAC 780
Db 721 ACTCAGCTGAGCAACATGACGGATCCGAAAGACGGTTCCGCTGTGTTGTCGGGCTACACCGAC 780

Qy 781 CGCATCGGTTCCGAAGCTTACAAAGCGTTCAGTGTCTGAGAAAGTGCTCAGTCCGTGTT 840
Db 781 CGCATCGGTTCCGAAGCTTACAAAGCGTTCAGTGTCTGAGAAAGTGCTCAGTCCGTGTT 840

Qy 841 GACTACCCGGTTGCTAAAGGCATCCCGGCTCGCAAATCTCCGCGCTCCGCGCTCGCGCTCGCGCTGAA 900
Db 841 GACTACCCGGTTGCTAAAGGCATCCCGGCTCGCAAATCTCCGCGCTCCGCGCTCGCGCTCGCGCTGAA 900

Qy 901 TCCAACCCGGTTACTGCAACACCTGACAACGTGAAAGCTGAAACGTGATCGAT 960
Db 901 TCCAACCCGGTTACTGCAACACCTGACAACGTGAAAGCTGAAACGTGATCGAT 960

Qy 961 TGCCTGGCTCCGGATCCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
Db 961 TGCCTGGCTCCGGATCCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy 1021 CAGCCCGCGCGGGTTAA 1035
Db 1021 CAGCCCGCGCGGGTTAA 1035

Query Match   70.0%; Score 724.4; DB 9; Length 730;
Best Local Similarity 99.9%; Pred. No. 5.4e-226;
Matches 725; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

```
Qy  276 GCTGACCGCTAAACTGGGTTACCCGATCACTGACGATCTGGACATCTACACCCCGCTGGG 335
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    1 GCTGACCGCTAAACTGGGTTACCCGATCACTGACGATCTGGACATCTACACCCCGCTGGG 60

Qy  336 CGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTTCCCGTAG 395
         |||||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db   61 CGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCAACCGGCGTTCCCGTAG 120

Qy  396 CGAACACGACACTGGCGTTCCCCAGTATTGCTGGCGGCGTAGAGTGGCTGTTACTCG 455
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CGAACACGACACTGGCGTTCCCCAGTATTGCTGGCGGCGTAGAGTGGCTGTTACTCG 180

Qy  456 TGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCGGGCACTGT 515
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 TGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCGGGCACTGT 240

Qy  516 GGGTACCCGCTCTGATAACGGCATGCTGAGCCTGGCGTTCCTACCGGCGTTCGGTCAGGA 575
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGGTACCCGCTCTGATAACGGCATGCTGAGCCTGGCGTTCCTACCGGCGTTCGGTCAGGA 300

Qy  576 AGATGCTGCACCCGTTGTTGCTCCGGCGCATGGTGCTGCTACCAAGCA 635
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AGATGCTGCACCCGTTGTTGCTCCGGCGCATGGTGCTGCTACCAAGCA 360

Qy  636 CTTCACCCGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCGGAAGG 695
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CTTCACCCGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCGGAAGG 420

Qy  696 TCAGCAGGCTCTGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAAGACGGTTC 755
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TCAGCAGGCTCTGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAAGACGGTTC 480

Qy  756 CGCTGTTGTTCTGGGCTACACCACCGACCGATCGGTTCTACAACCAGCAGCTGTC 815
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CGCTGTTGTTCTGGGCTACACCACCGACCGATCGGTTCTACAACCAGCAGCTGTC 540

Qy  816 TGAGAAACGTGCCAGTCCGTTGTGACTACCGGTTGCTAAAGGCATCCCGGCTGGCAA 875
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 TGAGAAACGTGCCAGTCCGTTGTGACTACCGGTTGCTAAAGGCATCCCGGCTGGCAA 600

Qy  876 AATCTCCGCTCGCGGCATGGGTGAATCCAACCCGTTACTGGCAACACGT 935
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AATCTCCGCTCGCGGCATGGGTGAATCCAACCCGTTACTGGCAACACGT 660

Qy  936 GAAAGTTCGGCTGCCCTGATCGATTGCGTTCCGATCGTCGTGTAGAGATCGAAGT 995
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 GAAAGTTCGGCTGCCCTGATCGATTGCGTTCCGATCGTCGTGTAGAGATCGAAGT 720

Qy  996 TAAAGG 1001
         ||||||
Db  721 TAAAGG 726
```

RESULT 4
US-10-121-120-11

```
Db  181 ...|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        GACTGGCTGGGCCGTATGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG 240

Qy  268 GGCGGTTCAGCTGACCGCTAAACTGGGTTACCCGATCACTGACGATCTGGACATCTACACC 327
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GGGGTTCAGCTGACCGCTAAACTGGGTTACCCGATCACTGACGATCTGGACATCTACACC 300

Qy  328 CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT 387
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT 360

Qy  388 TCCCGTAGCGAACACGACACTGGCGTTCCCCAGTATTGCTGGCGGCGTAGAGTGGCT 447
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TCCCGTAGCGAACACGACACTGGCGTTCCCCAGTATTGCTGGCGGCGTAGAGTGGCT 420

Qy  448 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 507
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 480

Qy  508 GGCACTGTGGGTACCCGCTCTGATAACGGCATGCTGAGCCTGGCGTTCCTACCGGCGTTC 567
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GGCACTGTGGGTACCCGCTCTGATAACGGCATGCTGAGCCTGGCGTTCCTACCGGCGTTC 540

Qy  568 GGTCAGGAAGATGCTGCACCCGTTGTTGCTCCGGCGCATGGTGCTGCTACCAAGAGTGGCT 627
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GGTCAGGAAGATGCTGCACCCGTTGTTGCTCCGGCGCATGGTGCTGCTACCAAGAGTGGCT 600

Qy  628 ACCAAGCACTTCACCCGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA 687
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ACCAAGCACTTCACCCGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA 660

Qy  688 CCGGAAGGTCAGCAGGCTCTGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 747
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 CCGGAAGGTCAGCAGGCTCTGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 720

Qy  748 GACGGTTCCGCTGTTGTTCTGGGCTACACCACCGACCGATCGGTTCTACAACCAG 807
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GACGGTTCCGCTGTTGTTCTGGGCTACACCACCGACCGATCGGTTCTACAACCAG 780

Qy  808 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 867
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 840

Qy  868 GCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGTTACTGGCAACACCTGT 927
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGTTACTGGCAACACCTGT 900

Qy  928 GACAACGTGAAAGCTGCGCGTGCCCTGATCGATTGCGTTCCGGCTGCTCCGGATCGTCGTGTAGAG 987
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GACAACGTGAAAGCTGCGCGTGCCCTGATCGATTGCGTTCCGGCTGCTCCGGATCGTCGTGTAGAG 960

Qy  988 ATCGAAGTTAAAGGCTACAAAGAAGAAGTTGTAACTCAGCCGCCGCGGGTTAA 1035
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 ATCGAAGTTAAAGGCTACAAAGAAGAAGTTGTAACTCAGCCGCCGCGGGGTTAA 1008
```

RESULT 3
US-09-452-599-11

```
; APPLICANT: Ouellette, Marc
; APPLICANT: Roy, Paul H.
; TITLE OF INVENTION: Specific and Universal Probes and Amplification
; TITLE OF INVENTION: Primers
; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical
; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
; FILE REFERENCE: 12287.31
; CURRENT APPLICATION NUMBER: US/10/121,120
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/452,599
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 08/304,732
; PRIOR FILING DATE: 1994-09-12
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 730
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-10-121-120-11

Query Match        70.0%;     Score 724.4;  DB 15;   Length 730;
Best Local Similarity  99.9%;  Pred. No. 5.4e-226;
Matches  725;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  0;

Qy    276  GCTGACCGCTAAACTGGGTTACCCGAGACATCTGGACATCTACACCCGTCTGGG  335
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GCTGACCGCTAAACTGGGTTACCCGAGACATCTGGACATCTACACCCGTCTGGG   60

Qy    336  CGGCATGGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCTTTCCCGTAG  395
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CGGCATGGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCAACCGGCTTTCCCGTAG  120

Qy    396  CGAACACGACACTGGCGTTTCCCCAGTATTTGCTGCGCGGCTAGAGTGGGCTGTTACTCG  455
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CGAACACGACACTGGCGTTTCCCCAGTATTTGCTGCGCGGCTAGAGTGGGCTGTTACTCG  180

Qy    456  TGACATCGCTACCCGTCTGAATACCAGTGGTTAACACATCGGCACGCGACGGGCACTGT  515
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGACATCGCTACCCGTCTGAATACCAGTGGTTAACACATCGGCACGCGACGGGCACTGT  240

Qy    516  GGGTACCCGTCTGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGA  575
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGGTACCCGTCTGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGA  300

Qy    576  AGATGCTGCACCGGTTGTTGCTGAGCGTTCGGGCTCCGGAAGTGGCTACCAAGCA  635
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AGATGCTGCACCGGTTGTTGCTGAGCGTTCGGGCTCCGGAAGTGGCTACCAAGCA  360

Qy    636  CTTCACCCTGAAGTCTGACGTTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAACCGGAAGG  695
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  CTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAACCGGAAGG  420

Qy    696  TCAGCAGGCTCGGATCAGCTGTACACTCAGCTGAGCAACATGGAACTCGGAAAGACGGTTC  755
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  TCAGCAGGCTCGGATCAGCTGTACACTCAGCTGAGCAACATGGAACTCGGAAAGACGGTTC  480

Qy    756  CGCTGTTGTTCTCGGGCTACACCGACCGGCATCGGTTCGAAGCTTACAACCAGCAGCTGTC  815
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CGCTGTTGTTCTCGGGCTACACCGACCGGCATCGGTTCGAAGCTTACAACCAGCAGCTGTC  540

Qy    816  TGAGAAACGTGCTCAGTCCGGTTGCTAAAGGCATCCGGCTTGGCAA  875
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TGAGAAACGTGCTCAGTCCGGTTGCTAAAGGCATCCGGCTTGGCAA  600

Qy    876  AATCTTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGTGACAACGT  935
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  AATCTTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGTGACAACGT  660

Qy    936  GAAAGCTCGCGCTGTCTCGCGGGATCGGTTCGTGTAGAGATCGAAGT  995
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GAAAGCTCGCGCTGTCTCGCGGGATCGGTTCGTGTAGAGATCGAAGT  720

Qy    996  TAAAGG  1001
           ||||||
Db    721  TAAAGG   726
```

```
RESULT 5
US-10-425-821-69
; Sequence 69, Application US/10425821
; Publication No. US20040219530A1
; GENERAL INFORMATION:
; APPLICANT: BROUSSEAU, Roland
; APPLICANT: HAREL, Jos,e
; APPLICANT: BEKAL, Sadjia
; TITLE OF INVENTION: ARRAY AND USES THEREOF
; FILE REFERENCE: 86369-3
; CURRENT APPLICATION NUMBER: US/10/425,821
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
;   LENGTH: 1422
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-425-821-69

Query Match        67.3%;     Score 696.2;  DB 18;   Length 1422;
Best Local Similarity  82.1%;  Pred. No. 1.2e-216;
Matches  838;  Conservative  0;  Mismatches  153;  Indels  30;  Gaps  2;

Qy     15  CGTACTGAATGCGGCCTCCGAAAGATAACACACTGGTATGCAGGTGGTAAACTGGGTTGGTC   74
           |||||| |||||||| |||||||||||||||||||||||| |||||| ||||||||| |||
Db    337  CGTAGCGCAGGCCGCGTCCGAAAGATAACACACTGGTACACTGGTCTAAACTGGGCTGTC  396

Qy     75  CCAGTATCACGACACCCGTTTCTACGGTAACCGTTCAGAACACAACGGTCCGACCCG  134
           |||||| |||||  |||||||           |||||| |||||||||||||||
Db    397  CCAGTACCATGATACTGGTTTCATC--------AACAACAATGGCCCGACCCA  441

Qy    135  TAACGATCAGCTTGGTGTGCGTTCGGTGGTTACCAGGTTAACCGTAGTTGGCGTT  194
           |||||  |||| |||||||| |||| ||||||||||||||||||||| |||||||
Db    442  TGAAAACCAACTGGGCGGCCGTGCTTTTGGTGGTTACCAGGTTAACCGTATGTTGGCTT  501

Qy    195  CGAAATGGGTTATGACGTTCAGCTGACGTGGCGTGCATATAAAGGCAGCGTTGCACACGGTGC  254
           |||| |||||||||||||||||| |||| |||| ||||||||||||||||| ||||||||||
Db    502  TGAAAATGGGTTACGACGGTTAGGTGTATGCGTACAAAAGCGGTTGAAAACGGTGC  561

Qy    255  TTTCAAAGCTCAGGCGGTTCAGCTGACGTTACCCGGTTACCGCTAAACTGGGTTACCGATCACTGACGATCT  314
           ||||||||||||||||||||||||||| |||||||||||||| |||||||||||||||||||
Db    562  ATACAAAGCTCAGGGGCGTTCAACTGACGCGTTACCCGGTTACCGATCACTGACGACCT  621

Qy    315  GGACATCTACACCCGTCTGGGCGGCCATGGTTTGGCGCGTCACGGCGAACTACGC  374
           |||||||||||||||||||||||||| ||||||| |||||| |||||||||||||
Db    622  GGACATCTACACTCGTCTGGGTGGCATGGATGGTGCAGCACACTAAATCCAA------  675

Qy    375  TTCTACCGGCGTTTCCCGTAGGAGCAACAACACGACACCGGAAAGACGGTTC  434
           |||| ||||||||||||| | |||||||||| ||||||||| ||||||||||
Db    676  --------CGTTTATGGTAAAAACCACAGACACCGCGTTTCTCCGGCTGCGCGG  726

Qy    435  CGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGTTAACAA  494
           |||| ||||||||||||||| |||||| |||||||||||||||| |||||||||||||
Db    727  TGTTGAGTGACGGGATCACTCCTGAAATCGCTACCCGTCTGGAATACCAGTGACGACCAA  786

Qy    495  CATCGGCGACGCGGGGACACTGGTTGGTTACCGCATGCTGGAGCCTGGGCGGT  554
           ||||| |||||| ||||||||||| ||||| ||||||||||||||||||||||
Db    787  CATCGGTGACGCAGCACACCATCGGCTCGTTCCGGAGCATGCTGAGCCTGGGCGGTGT  846

Qy    555  TTCCTACCGCTTCGGGTACCGGTTCAGGAGGCCTCCAGTAGTTGCTCCGGCTCCGGCCGC  614
           |||| |||||||||||||||||||||| |||||||| ||||||| ||||||||||||||
Db    847  TTCCTACCGTTTCGGTTACCGGGCGAGGCGAGGCAGCTCCAGTAGTTGCTCGGCTCCGGC  906

Qy    615  TCCGGAAGTGGCTACCAAGACACTCACCCTGAAGTCTGAAGTTCTGTTCAACTTCAACAA  674
           |||||||||| |||| ||||||||||||||||||||||||| ||||| ||||||||||||
Db    907  ACCGGAAGTACAGACACCAGCACTTCACCCTGAAGTCTGAAGATCTGTTCTCAACTTCAACAA  966
```

Qy   675 AGCTACCCTGAAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAA 734
         ||||||||||||||||||||||||||||||||||| ||||| |||||||||
Db   967 AGCAACCCTGAAAACCGGAAGGTCAGCAGGCTCTCCTCGGATCAGCCAGCTGAGCAA 1026

Qy   735 CATGGATCCGAAAGACAGCGGTCCGCTGTTGTTCTGGCGTACACCGACCGACCGGTTCCGA 794
         |||||||||||||||||||||| ||| |||||||||||| |||||||||||| ||||
Db 1027 CTTGGATCCGAAAGACAGCGGTTCCGTAGTTGTTCTGGGTTACACCGACCGGTTTCTGA 1086

Qy   795 AGCTTACAACCAGCAGCTGCTGTCTGAGAAACGTGCTCAGTCCGTTGACTACCTGGTTGC 854
         |||||||||||||||| || ||| |||||||||| ||||| |||||||||||||
Db 1087 CGCTTACAACCAGCAGGGTCGTGTCCGAGCGGCGGCGCTCGTGTTGTTGTTACCTGATCTC 1146

Qy   855 TAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTAC 914
         |||||| |||||||| || ||||||||||| ||| |||||||||||||||||||||||||
Db 1147 CAAAGGTATCCCGGCAGACAAGAATCTCCCACGTGGTATGGGCGAATCCAACCCGGTTAC 1206

Qy   915 TGGCAACCTGTGACAACGTGAAAGCTCGGCTCGCCCTGATCGATTGCCTGCTCCGGA 974
         |||||| |||||||||||||||| |||||||| |||||||||||||||||||||||
Db 1207 TGGCAACACCTGTGACAACGTGAAACAGCGTGCTGCACTGACGATCGATCGCTGCCTGCTCCGGA 1266

Qy   975 TCGTCGTGTAGAGATCGAAGTTAAAGCTACAAAGAGTTGTAACTCAGCCGGCGGGGTTA 1034
         ||||| |||||||||||||| ||| |||||||||||||||||||||||||||| |||||
Db 1267 TCGTCGGCGTAGAGATCGAAGTATCAAAGGTATCAAAGAGACGTTGTAACTCAGCCGCAGGCTTA 1326

Qy 1035 A 1035
         |
Db 1327 A 1327

RESULT 6
US-10-432-056-1
; Sequence 1, Application US/10432056
; Publication No. US20040014661A1
; GENERAL INFORMATION:
; APPLICANT: GOETSCH, Liliane
; APPLICANT: HAEUW, Jean-Francois
; APPLICANT: ROBERT, Alain
; TITLE OF INVENTION: PERIPLASMIC DOMAIN OF AN ENTEROBACTERIUM OMP PROTEIN
; TITLE OF INVENTION: AND ITS USE AS CARRIER OR ADJUVANT
; FILE REFERENCE: 344 667 - US
; CURRENT APPLICATION NUMBER: US/10/432,056
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 596
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: FR 00/14 909
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
;
SEQ ID NO 1
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(411)
US-10-432-056-1

Query Match   38.9%; Score 403; DB 16; Length 411;
Best Local Similarity 98.8%; Pred. No. 5.8e-121;
Matches 406; Conservative 0; Mismatches 0; Indels 5; Gaps 0;

Qy   622 GTGGGCTACCAAGCACTTCACCCTGAAGCTCGACGTTCGTTCTGTTGTTCAACTTCAACAAAGCTACC 681
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTGGGCTACCAAGCACTTCACCCTGAAGCTCTGACGTTCGTTCTGTTGTTCAACTTCAACAAAGCTACC 60

Qy   682 CTGAAACCGGAAGGTCAGCAGGCTCAGGCTCTCGATCAGCTGTACACTCAGCTGAGCAACATGGAT 741
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGAAACCGGAAGGTCAGCAGGCTCAGGCTCTCGATCAGCTGTACACTCAGCTGAGCAACATGGAT 120

Qy   742 CCGAAAGACAGCGGTTCCGCTGTTGTTCTGGCGTACACCGACCGGCTACACCGGTTCCGAAGCTTAC 801
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCGAAAGACAGCGGTTCCGCTGTTGTTCTGGCGTACACCGACCGGCTACACCGGTTCCGAAGCTTAC 180

RESULT 7
US-10-416-708A-23
; Sequence 23, Application US/10416708A
; Publication No. US20040161753A1
GENERAL INFORMATION:
; APPLICANT: Wise, John G.
; APPLICANT: Fromknecht, Katja
; TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
; TITLE OF INVENTION: SPECIFICITIES
; FILE REFERRNCE: 37779-0004
; CURRENT APPLICATION NUMBER: US/10/416,708A
; CURRENT FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.2
;
SEQ ID NO 23
; LENGTH: 3630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid derived from pComp containing a hexahisadine separator
; OTHER INFORMATION: gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(600)
; OTHER INFORMATION: pComp derived DNA
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (6)..(23)
; OTHER INFORMATION: target binding sequence from cauliflower mosaic virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)..(624)
; OTHER INFORMATION: ompA his tag fusion
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (601)..(624)
; OTHER INFORMATION: linker and histag
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (658)..(3630)
; OTHER INFORMATION: pComp derived DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (688)..(1137)
; OTHER INFORMATION: lacZ' fragment
US-10-416-708A-23

Query Match   29.3%; Score 303.4; DB 17; Length 3630;
Best Local Similarity 74.3%; Pred. No. 5.1e-88;
Matches 423; Conservative 0; Mismatches 116; Indels 30; Gaps 2;

Qy   15 CGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGTAAACTGGGTTGGTC 74
         |||| | |||||||||||||||||||||||||||| |||||| ||||||||| ||||
Db   111 CGTAGGCGCAGGCGCCTCCGAAAGATAACACCTGGTACACCTGGTGTGTCTAAACTGGGTGGTC 170

Qy   802 AACCAGCAGCTGCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGC 861
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AACCAGCAGCTGCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGC 240

Qy   862 ATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGCGCAAC 921
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTACTGCGCAAC 300

Qy   922 ACCTGTGACAACGTGAAAGCTCGGCTCGCCCTGCCCTGATTGCTTGCCTGCTCCGATCGTCGT 981
         ||||||||||||||||| ||||||||||||||| |||||||||||||||||||||||||
Db   301 ACCTGTGACAACGTGAAAGCTCGGCTCGCCCTGCCCTGATTGCTTGCCTGCTCCGATCGTCGT 360

Qy   982 GTAGAGATCGAAGTTAAAGCTACAAAGAAGTTGTAACTCAGCCCGCGGGT 1032
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCCTCAGGCT 411

```
Qy    75  CCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCG  134
              ||||||| || ||||| || ||||||                    ||||||| || ||||
Db   171  CCAGTACCCATGACACTGGTTTCATC-----------------AACAACAATGGCCGACCCA  215

Qy   135  TAACGATCAGCTTGGTGCTGGTGCCGTATATAAAGGCAGCGTTGACAACGGTGC  194
              |  || ||||||  ||||||| ||||| ||||||||||||| | |||||||
Db   216  TGAAAACCAACTGGGGCGCTGTGCTTTTGGTGGTTACCAGTTAACCGTATGTTGGCTT  275

Qy   195  CGAAATGGGTTATGACTGGCTGGGCCGTATGTGCATATAAAGGCAGCGTTGACAACGGTGC  254
              ||||||||||||| || |||||| ||||||||||| ||||||||||||||||||||||||
Db   276  TGAAATGGGTTACGACTGGTGTAGGTGTCGTATGCCGTACAAAGGCAGCGTTGAAAACGGTGC  335

Qy   255  TTTCAAAGCTTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCGTATGTTGACCGTTAACAA  314
              ||||||||||||||||||| |||| ||||||||||| ||||| || ||||||||||| |||
Db   336  ATACAAAGCTTCAGGGCGTTCAACTGACCGCTAAACTGGGTTACCCAATCACTGACGACCT  395

Qy   315  GGACATCTACACCCGTCTGGGGCGCCATGGTTGGCGCGTGATGGCATGGCAACTACGC  374
              |||||| |||| || |||| ||| | ||  ||| ||| || |||  ||||||||
Db   396  GGACATCTACACTCGTCTGGGTGGTGGTGGCATGGTGGCGTGCAGACACTAAATCCAA------  449

Qy   375  TTCTACCGGCGTTTCCCGTTTCTCCGGTCTTCGCTGGCGG  434
              ||||| ||||| || |||  || ||  |||| ||||| ||||
Db   450  CGTTTATGGTAAAAACCACGACACCCGCGTTCGCTGCGCGG  500

Qy   435  CGTAGAGTGGGCTGTTACTCGTGACATCGCTAAAATCGCTAACTGGT  494
              |||| ||||| ||||  || | ||| || |||||| ||  ||||||
Db   501  TGTTGAGTACGCGATCACTCCTGAAATCGCTGAAATCGCTGAAATCGCT  560

Qy   495  CATCGGCGACGCGGGCACTGTGGGTACCCGTCCGATAACGGCGT  554
              |||||||||||||| || ||| ||||| || ||||| ||||
Db   561  CATCGGTGACGCGGGCACTGTGGGTACCCGTCCGATAACCGC  603
```

RESULT 8
US-10-416-708A-26
; Sequence 26, Application US/10416708A
; Publication No. US2004016175A1
; GENERAL INFORMATION:
; APPLICANT: Wise, John G.
; APPLICANT: Fromknecht, Katja
; TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
; TITLE OF INVENTION: SPECIFICITIES
; FILE REFERENCE: 37779-0004
; CURRENT APPLICATION NUMBER: US/10/416,708A
; CURRENT FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 26
; LENGTH: 3603
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid derived from pComp containing a flag tag separator gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(600)
; OTHER INFORMATION: pComp derived DNA
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (6)..(23)
; OTHER INFORMATION: target binding site
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)..(630)
; OTHER INFORMATION: ompA flag tag fusion
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600)..(630)
; OTHER INFORMATION: linker flag tag

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (631)..(3603)
; OTHER INFORMATION: pComp derived DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (661)..(1110)
; OTHER INFORMATION: lacZ' fragment
US-10-416-708A-26

```
Query Match         29.0%;     Score 300.6; DB 17;  Length 3603;
Best Local Similarity  77.2%;   Pred. No. 4.2e-87;
Matches 404; Conservative   0; Mismatches  89; Indels  30; Gaps  2;

Qy    15  CGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTGGTC  74
              ||  || ||| | ||||| || || ||||||| ||| |||||||  |||||| ||||
Db   111  CGTAGCGCAGGCGCCGTCCGAAAGATAACACCTGGTACACTGGGTGCTAAACTGGGCTCGTC  170

Qy    75  CCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCG  134
              |||||| || ||||| || |||                    |||||| || ||||
Db   171  CCAGTACCCATGACACTGGTTTCATC-----------------AACAACAATGGCCGACCCA  215

Qy   135  TAACGATCAGCTTGGTGCTGGTGCCGTATATAAAGGCAGCGTTGACAACGGTGC  194
              |  || ||||||  ||||||| ||||| |||||||||||| | |||||||
Db   216  TGAAAACCAACTGGGGCGCTGTGCTTTTGGTGGTTACCAGTTAACCGTATGTTGGCTT  275

Qy   195  CGAAATGGGTTATGACTGGCTGGGCCGTATGTGCATATAAAGGCAGCGTTGACAACGGTGC  254
              ||||||||||||| || |||||| ||||||||||| ||||||||||||||||||||||||
Db   276  TGAAATGGGTTACGACTGGTGTAGGTGTCGTATGCCGTACAAAGGCAGCGTTGAAAACGGTGC  335

Qy   255  TTTCAAAGCTTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCGTATGTTGACCGATCT  314
              ||||||||||||||||||| |||| ||||||||||| ||||||| |||||| |||||||
Db   336  ATACAAAGCTTCAGGGCGTTCAACTGACCGCTAAACTGGGTTACCCAATCACTGACGACCT  395

Qy   315  GGACATCTACACCCGTCTGGGGCGCCATGGTTGGCGCGTGATGGCATGGCAACTACGC  374
              |||||| |||| || |||| ||| | ||  |||  ||| || |||  ||||||||
Db   396  GGACATCTACACTCGTCTGGGTGGTGGTGGCATGGTGGCGTGCAGACACTAAATCCAA------  449

Qy   375  TTCTACCGGCGTTTCCCGTTTCTCCGGTCTTCGCTGGCGG  434
              ||||| ||||| || |||  || ||  |||| ||||| ||||
Db   450  --------CGTTTATGGTAAAAACCACGACACCCGCGTTCGCTGCGCGG  500

Qy   435  CGTAGAGTGGGCTGTTACTCGTGACATCGCTAAAATCGCTAACTGGTGGTTAACAA  494
              |||| ||||| ||||  || | ||| || |||||| ||  |||||| ||
Db   501  TGTTGAGTACGCGATCACTCCTGAAATCGCTGAAATCGCTGAAATCGCTGAAATCGCT  560

Qy   495  CATCGGCGACGCGGGCACTGTGGGTACCCGTCCGATAACGGC  537
              |||||||||||||| || ||| ||||| || ||||| ||||
Db   561  CATCGGTGACGCGGGCACTGTGGGTACCCGTCCGGACAACGGC  603
```

RESULT 9
US-10-416-708A-9
; Sequence 9, Application US/10416708A
; Publication No. US2004016175A1
; GENERAL INFORMATION:
; APPLICANT: Wise, John G.
; APPLICANT: Fromknecht, Katja
; TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
; TITLE OF INVENTION: SPECIFICITIES
; FILE REFERENCE: 37779-0004
; CURRENT APPLICATION NUMBER: US/10/416,708A
; CURRENT FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 9
; LENGTH: 3660
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid containing a promoter, cauliflower mosaic virus-derived
; OTHER INFORMATION: target operator and reporter gene and ompA-derived separator gene
; FEATURE:

; NAME/KEY: misc_binding
; LOCATION: (1)..(28)
; OTHER INFORMATION: cauliflower mosaic virus derived target DNA sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)..(636)
; OTHER INFORMATION: ompA separation tag fusion
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (64)..(608)
; OTHER INFORMATION: E. coli ompA derived DNA sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (608)..(684)
; OTHER INFORMATION: separation tag DNA sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (685)..(3660)
; OTHER INFORMATION: pP2RHIV1-derived DNA sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (718)..(1164)
; OTHER INFORMATION: lacZ' fragment
US-10-416-708A-9

Query Match   28.9%; Score 298.6; DB 17; Length 3660;
Best Local Similarity   77.2%; Pred. No. 1.9e-86;
Matches   402; Conservative   0; Mismatches   89; Indels   30; Gaps   2;

Qy   15 CGTACTGAATGGGCGGTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTGGTC 74
          ||||| | |||| |||||||||||||||||||||||||| ||||||||||||| |||||
Db   111 CGTAGGCGCAGGCCGCTCCGAAAGATAACACCTGGTCACTGGTGCTAAACTGGGCTGGTC 170

Qy   75 CCAGTATCACGACACCGGTTTCACGGTAACGGTTCCAGAACAACAACGGTCCGACCG 134
          |||||| ||||||| ||||| ||||||||||| |||||||
Db   171 CCAGTACCATGACACTGGTTTCATC--------------AACAACAATGGCCCGACCCA 215

Qy   135 TAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTACCCGTACCTCGGTTT 194
          | |||||||||| ||| ||| ||| |||| |||||||||||||| ||||| ||||||
Db   216 TGAAAACCAACTGGGCGCTGGTGCTTTTGGTGTTGGTGTTGGTTACCAGGTTACCCGTATGTTGGCTT 275

Qy   195 CGAAAATGGGTTATGACTGGCTCGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGC 254
          ||||||||| |||||||||| | || |||||| |||| || |||||||||||||||||||
Db   276 TGAAAATGGGTTACGACTGGTTAGGTCGTGATGCCGTACAAAGGCAGCGTTGAAAAACCGGTGC 335

Qy   255 TTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGTTACCCGTCCTGATAACG 535
          ||||||||||||||||||||||| ||||||||||||||||||||| || ||||||
Db   336 ATACAAAGCTCAGGGCGTTCAACTGACCGCTAAACTGGTTACCCAATCACTCACTGACGACCT 395

Qy   315 GGACATCTACACCCGTCTGGGCGGCACGTGGTTGGCGCGCTGACTCCAAAGGCAACTACGC 374
          |||||||||||| ||||| |||| |||||||| ||||||||||| | |||||
Db   396 GGACATCTACACTCGTCTGTGGGCATGGTGGTGCGTGTGCAGACACTAAATCCAA------ 449

Qy   375 TTCTACCGGGCGGTTTCCCGTAGCGGACACCTGGCGTTCTCCGGTTTTCTCGGTTCGGCGG 434
          ||||| |||||||||||||||||| |||| | ||||| | |||| | |||||
Db   450 ---------CGTTTATGGTAAAAACCACGACACCGGTGATTATGGTGGCGTTCGGCGG 500

Qy   435 CGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTCGGAATACCAGTGGTTAACAA 494
          ||||| | ||||| | |||| ||||||| ||||||||| ||||||||||| ||||
Db   501 TGTTGAGTACGGCGGATCACTCCTCGAAATCGCTACCCGTCTCCTGAATACCAGTGGACCAACAA 560

Qy   495 CATCGGCGCGCGGGCACTGTGGTCACCGTCCGTCCGGACAACG 535
          |||||| | ||| ||||||| |||||||| |||||||||||
Db   561 CATCGGTGACGCACACACCATCGGCACTCGTCCGGACAACG 601

RESULT 10
US-10-416-708A-62
; Sequence 62, Application US/10416708A
; Publication No. US2004161753A1
; GENERAL INFORMATION:
; APPLICANT: Wise, John G.
; APPLICANT: Fromknecht, Katja

; TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
; TITLE OF INVENTION: SPECIFICITIES
; FILE REFERENCE: 37779-0004
; CURRENT APPLICATION NUMBER: US/10/416,708A
; CURRENT FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; LENGTH: 4543
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid 2 for bacterial transcriptional activation screening
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (12)..(63)
; OTHER INFORMATION: zif binding site and spacer
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (68)..(110)
; OTHER INFORMATION: weak promoter
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(799)
; OTHER INFORMATION: HIS3 fragment
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (800)..(847)
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: (3068)..(3119)
; OTHER INFORMATION: zif binding and spacer
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (3124)..(3166)
; OTHER INFORMATION: weak promoter
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3198)..(3782)
; OTHER INFORMATION: ompA separation tag fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3861)..(4310)
; OTHER INFORMATION: lacZ' fragment
US-10-416-708A-62

Query Match   28.9%; Score 298.6; DB 17; Length 4543;
Best Local Similarity   77.2%; Pred. No. 2.1e-86;
Matches   402; Conservative   0; Mismatches   89; Indels   30; Gaps   2;

Qy   15 CGTACTGAATGGGCGGTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTGGTC 74
          ||||| | |||| |||||||||||||||||||||||||| ||||||||||||| |||||
Db   3254 CGTAGGCGCAGGCCGCTCCGAAAGATAACACCTGGTCACTGGTGCTAAACTGGGCTGGTC 3313

Qy   75 CCAGTATCACGACACCGGTTTCACGGTAACGGTTCCAGAACAACAACGGTCCGACCG 134
          |||||| ||||||| ||||| ||||||||||| |||||||
Db   3314 CCAGTACCATGACACTGGTTTCATC--------------AACAACAATGGCCCGACCCA 3358

Qy   135 TAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCGTACCTCGGTTT 194
          | |||||||||| ||| ||| ||| |||| |||||||||| ||||| ||||| ||||||
Db   3359 TGAAAACCAACTGGGCGCTGGTGCTTTTGGTGTTGGTGTTGGTTACCAGGTTACCCGTATGTTGGCTT 3418

Qy   195 CGAAAATGGGTTATGACTGGCTCGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGC 254
          ||||||||| |||||||||| | || |||||| |||| || |||||||||||||||||||
Db   3419 TGAAAATGGGTTACGACTGGTTAGGTCGTGATGCCGTACAAAGGCAGCGTTGAAAAACCGGTGC 3478

Qy   255 TTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGTTACCCGTGATCACTGACGATCT 314
          ||||||||||||||||||||||| ||||||||||||||||||||| || ||||| |||
Db   3479 ATACAAAGCTCAGGGCGTTCAACTGACCGCTAAACTGGTTACCCAATCACTCACTGACGACCT 3538

Qy   315 GGACATCTACACCCGTCTGGGCCGGCATGGTTTGGCGCGCTGACTCCACGCAACTACGC 374
          |||||||||||| ||||| | || ||||| | | | | ||| |||
Db   3539 GGACATCTACACTCGTCTCTGGTGGTGGCATGTGCGTGATGCGTGCGAGACACTAAATCCAA------ 3592

```
Qy   375  TTCTACCGGCGTTCCCGTAGCGAACACGACACTGGCGTTCCCCAGTATTGCTGGCGG  434
          |||    ||||  |||||||||||||||||||||||||| ||    |  ||||||||
Db  3593  --------CGTTTATGGTAAAAACCACGACACCGGCGTTTCTCCGGTCTTCGCTGGCGG  3643

Qy   435  CGTAGAGTGGGCTGTTACTCGTGACATCGTACCCGTCTGGAATACCAGTGGGTTAACAA  494
          ||| || ||  || |  ||||||||||| | |||||||||||||||||||||| ||||||
Db  3644  TGTTGAGTACGACGATCACTCTGAAATCGCTGAATACCAGTGGGACCAACAA  3703

Qy   495  CATCGGCGACGCGGGCACTGTGGGCACTGTGGGTACCCGTCCTGATAACG  535
          |||||| ||||| ||| | ||| |||| | ||| |  ||||||| ||  |
Db  3704  CATCGGTGACGCACACACCATCGGCACTCGTCCGGACAAACG  3744
```

RESULT 11
US-10-033-399B-24
; Sequence 24, Application US/10033399B
; Publication No. US2003010455A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Caili
;   APPLICANT: Zhong, Pingyu
;   APPLICANT: Wang, Xinwei
;   TITLE OF INVENTION: ADAPTER-DIRECTED DISPLAY SYSTEMS
;   FILE REFERENCE: 13403.0005NPUS00
;   CURRENT APPLICATION NUMBER: US/10/033,399B
;   CURRENT FILING DATE: 2002-11-15
;   NUMBER OF SEQ ID NOS: 26
;   SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 24
;   LENGTH: 3817
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic, comprising GABAB receptor 2, Lpp-OmpA gene sequence
US-10-033-399B-24

Query Match   21.2%;   Score 219; DB 15;   Length 3817;
Best Local Similarity   76.3%;   Pred. No. 2.1e-60;
Matches 289; Conservative   0; Mismatches 75; Indels 15; Gaps 1;

```
Qy   172  CAGGTTAACCCGTACCTCGTTTCGAAATGGGTTATGACTGGGCCGTATGGCATAT  231
          || |||||||||||  || |||||||||||  ||||||||||  ||||| ||||||
Db   262  CTGGGCAACCCGTATGTTGGCTTTGAAATGGTTACGACTGGTTAGGTCGTATGCCGTAC  321

Qy   232  AAAGGCAGCGTTGACAAACGGTGCTTTCAGCTGACCGGCCGTAAACTG  291
          ||||||||||||||  ||||||||||  || |||||||||||||||||
Db   322  AAAGGCAGCGTTGAAAACGGTGCATACAAAGCTCAACTGACCGGCCGTAAACTG  381

Qy   292  GGTTACCGGATCAGCGATCTGGACATCTACACCGTCGGGCGGCCATGGTTTGGCGC  351
          |||||||  |||| ||| |||||||||||||| | |||||| ||||| ||||| ||
Db   382  GGTTACCCAATCACTGACGACCTGGACATCTACACTGCTGTGGGCGTGGTTATGGCGT  441

Qy   352  GCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACACGACACTGGC  411
          || ||| ||||||||||||| | |||  ||                |||||||||||||
Db   442  GCAGACACTAAATCCAA-----------CGTTTATGTGAAAAACCACGACACCGGC  486

Qy   412  GTTTCCCCAGTATTGCTGGCGGCTAGAGTGGGCTGTGTTACTACTCGTGACATCGCTACCCGT  471
          ||||| |  || ||||||||||  ||||| ||||||||| | ||||||||||||||||||||
Db   487  GTTTCTCCGGTCTTCGCTGGCGGTGTTGAGTACGACGATCACTCTGAAATCGCTACCCGT  546

Qy   472  CTGGAATACCAGTGGTTAACAACATCGGCGACGCGGGCACTGTGTGGCCGTCCTGAT  531
          |||||||||||||||  | |||||||| ||||| ||| || | |  ||||||| ||
Db   547  CTGGAATACCAGTGGACGACGACAACATCGGTGACGCACACACCATCCGCACTCGTCCGGAC  606

Qy   532  AACGGCATGCTGAGCCTGG  550
          |  ||  |  ||  ||||
Db   607  GGAGGTACATCCCGCCTGG  625
```

; GENERAL INFORMATION:
;   APPLICANT: Lowery E., David, et al.
;   TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
;   FILE REFERENCE: 28341/00435
;   CURRENT APPLICATION NUMBER: US/09/809,665A
;   CURRENT FILING DATE: 2001-03-15
;   PRIOR APPLICATION NUMBER: 60/153,453
;   PRIOR FILING DATE: 1999-09-10
;   PRIOR APPLICATION NUMBER: 60/128,689
;   PRIOR FILING DATE: 1999-04-09
;   PRIOR APPLICATION NUMBER: 09/545,199
;   PRIOR FILING DATE: 2000-04-06
;   NUMBER OF SEQ ID NOS: 197
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 150
;   LENGTH: 1095
;   TYPE: DNA
;   ORGANISM: Actinobacillus pleuropneumoniae
;   FEATURE:
;   OTHER INFORMATION: Omp5-2
;   NAME/KEY: CDS
;   LOCATION: (1)..(1092)
US-09-809-665A-150

Query Match   15.6%;   Score 161; DB 11;   Length 1095;
Best Local Similarity   56.4%;   Pred. No. 1.2e-41;
Matches 340; Conservative   0; Mismatches 260; Indels 3; Gaps 2;

```
Qy   414  TTCCCCAGTATTTGCTGGCGGCTAGAGTGGGCTGTTACTCGTGACATCGTACCCGTCT  473
          |||   | ||   || |||  ||  ||||  ||||  | ||| |  |  |||||||||
Db   483  TTCTTTAATTTTAGGTGCGGGTGTTGAGTACGACGAATTCTTCTGAATTAGCGGCACGTGT  542

Qy   474  GGAATACCAGTGGGTTAACACATCGGCGACGCGGCACTGTGGGTACCCGTCCTGATAA  533
          |||||||||| |||| ||||| |||  |||  ||| ||  |||||||| ||  |||||
Db   543  TGAATACCAATGGTTAAACAACGCAGGTAAAGCAACGCTACTCTACTTTAAATCGTATGGG  602

Qy   534  CGGCA-TGCTGAGCCTGGGCGTTTCTACCGCTTCGTCAGGAAGATGCTGC--ACCGGT  590
          |||||  |  |||  ||  ||  |  | |||||||||||  ||||  || |  ||||||
Db   603  TGCAACTGACTACCGTTCGGATATCAGTTCCGTATCGCAGGTTTAAGCTACCGGTTTCGG  662

Qy   591  TGTTGCTTCCGGCTCCGGCTCCGGGTTATGACTGGGCCGTACCCTCGAAGTC  650
          | ||||| |||||  ||||  ||| ||  || || ||  || | |||  |
Db   663  TCAAGGTGCGGCACCGGTGCTTGAGTACGACTGGTTAGGTGCAGTTGAAACTAAAAACTTCGCATTCAGCTC  722

Qy   651  TGACGTTCTGTTCAACTTCAACAAAGCTACCACCCCGCTCAGCGAGGTCAGCAGGCTCTGGA  710
          ||||||| | |||||||| ||||||| ||||  |||| |||| || |||||||| |||||||
Db   723  TGACGTATTATTCGCATTCGGAATCAAACTTAAAACCGGCTGCGCGCAACAGCATTAGA  782

Qy   711  TCAGCTGTACACTCAGCTGACGCTGAGCAACATGAGATCCGAAAGACGGTTCCCGTTGTTCTGGG  770
          || |  ||||  |||  ||||  |   |  ||| || |||  ||  ||||| | |  |  ||||
Db   783  TGCAATGCAAACCGAAATCAATAACGCAGGTTTATCAAATGCTGCAAGTAAACGG  842

Qy   771  CTACACCGACCGCATCGGTTCCGAACTTCCGAGCGTTCCCGTAGCGAACACGACACTGGC  830
          || |  || | |||  |  |  |||   | | || |||   | |||| ||  ||||||||
Db   843  TTACACGGACCGCATCCGTAAAGAAGCTTCAAACTTAAAACTTCACAAACGTCGTGCGGGA  902

Qy   831  GTTTCCCCAGTATTGCTGGCGGCTAGAGTTGCTAAAGGCATCCGCTACCCGT  890
          ||||| |  || ||||||||||  |  |  || || |||  || |||||||
Db   903  AACAGTAGCTAAACTACATCGTTTCTAAAGGTGCTCCGCAGCTAACGTAACGTCGTGTGAT  962

Qy   891  CATGGGTGAATCCAACCGGTTACTGCCACACCGGCACTGTGTGACAACCGTGAAAGCTCGCGCTGC  950
          || |||||||||   | ||||| ||  || | ||   | ||   ||   |  |||| ||||| ||
Db   963  TTACGGTGAAGCAAACCGGTTACCGGCGAACATGTGACAACATGTGACAAAGTTAAAGGTCGTAAAGC  1022

Qy   951  CCTGATCGATTGCCTGCTCCGGATCGTCGTGTAGAGATCGGAATCGAAGTTACACAAAGA  1010
          | |  |  ||  ||| |||  ||||  |  || || | || || |||||| ||||||||
Db  1023  ATTAATCGCTTAGCACCCGGATCGTCGTTGTTGAAGTTCAAGTTACTAAAGA  1082

Qy  1011  AGT  1013
          |||
Db  1083  AGT  1085
```

RESULT 12
US-09-809-665A-150
; Sequence 150, Application US/09809665A
; Publication No. US2004011026BA1

RESULT 13
US-10-854-299-150
; Sequence 150, Application US/10854299
; Publication No. US20050003512A1
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/10/854,299
; CURRENT FILING DATE: 2004-05-26
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 150
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Actinobacillus pleuropneumoniae
; FEATURE:
; OTHER INFORMATION: Omp5-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1092)
US-10-854-299-150

Query Match   15.6%;   Score 161;   DB 18;   Length 1095;
Best Local Similarity 56.4%;   Pred. No. 1.2e-41;
Matches 340;   Conservative 0;   Mismatches 260;   Indels 3;   Gaps 2;

```
Qy   414  TTCCCCAGTATTTGCTGGCGGCGTAGAGTGGGCTTACTGTGACATCGTTACCCGTCT  473
          ||| ||   | || |||  |  ||  || ||||  ||  || ||  || ||  ||
Db   483  TTCTTTAATTTTAGGTGCGGGGTGTTGAGTACGGCGCCAATTCTTCCTGAATTAGCGGCACGTGT  542

Qy   474  GGAATACCAGTGGGTTAACACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAA  533
          |||||||  |||| |||| ||| ||  | |  |||| ||| || || |||| | ||||
Db   543  TGAATACCCAATGGTTAAACAACCAGGTAAAGCAAGCTACTCACTTTAAATCGTATGGG  602

Qy   534  CGGCA-TGCTGAGCCTGGCGGTTTCCTACCGGTTCGGTCAGGAAGATGCTGC--ACCGGT  590
          ||||  ||||| | |||| | |  |  |||||| ||||  ||| |  || ||    ||||
Db   603  TGCAACTGACTACCGTTCGGATATCAGTTCGGTTTAAGCTACCGTTTCGG  662

Qy   591  TGTTGCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTC  650
          |  ||  | |||| ||| | ||| || |  | ||  |||| |||| ||| |||
Db   663  TCAAGGTGCGGCACCGGTTGCAGCTCCGGCGCCGGAGTTCCGGATATCAGTCGCAGGTTTCGG  722

Qy   651  TGACGTTCGTGTTCAACTTCAACAAAGGTACCCTGAAGGTCAGCAGGCTCTGGA  710
          ||| |||  |||| |||| ||  |||||  ||  || |  || |||| ||| |
Db   723  TGACGTATTATTCGCATTCGGTAAATCAAACTAAAAACTTCGCATTCAGCTC  782

Qy   711  TCAGCTGTACACTCAGCTGAGCAACAATGATCCGAAAGACGGTTCCGCTGTTGTTCTGGG  770
          ||||| ||| || |||| ||||  |||| ||  | |||| |||| |||||  || ||||
Db   783  TGCAATGCAAACCGAAATCAATAACGCAGGTTTATCAAAGTCGCGATCCAAGTAAACGG  842

Qy   771  CTACACCGACCGCATCGGTTCCGAAGCTTACACAACCAGCTCTGAGAAAACGTGCTCA  830
          |||||||| ||  ||||| | |||||| ||||| ||||| || ||||| | ||  |
Db   843  TTACACCGACCGTATCGGTAAAGAAGCTTCAAACTTTCACCACGTCGTGCGGA  902

Qy   831  GTCCGGTGTTGACTACCTGGTTGCTAAAGGCATCCCGCGCTGCGAAAATCTCCGGCTCGCGG  890
          || ||| || |||| ||| | |||| |||| ||| | |||  | |||| || |||| | |
Db   903  AACAGTAGCTAACTAACTGTTTCTAAAGGTGCTCCGCAGCTAACGTCAGTAGG  962

Qy   891  CATGGGGTGAATCCAACCGGTTACTGCGACACACCTGTGACAACGTGAAAGCTCGGCTGC  950
          |||||  ||| ||||  | |||||| | || || || |||||  || |||| |  ||
Db   963  TTACGGTGAAGCAAACCCGTTAACCGGCGCCGCAACATGTGACAAGTGACTAAAGC  1022

Qy   951  CCTGATCGATTGCCTGCCTCCGGCTCGTCGTGTAGAGATCGAAGTGAAGTTAAAGGCTACAAAGA  1010
          || |||| | ||| ||| || |||| |||| |  || ||| ||||  || ||||
Db  1023  ATTAATCGCTTGCTTAGCACCGGATCGTCGTTCGTTGTTGAAGTTCAAGTTACTACTAAAGA  1082
```

Qy   1011 AGT   1013
           |||
Db   1083 AGT   1085

RESULT 14
US-10-770-824-7
; Sequence 7, Application US/10770824
; Publication No. US20040198954A1
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Baarsch, Mary Jo
; APPLICANT: Rosey, Everett
; APPLICANT: Ankenbauer, Robert
; APPLICANT: Warren-Stewart, Lynn
; APPLICANT: Suiter, Brian
; APPLICANT: Keach, Robin
; TITLE OF INVENTION: NOVEL PROTEINS FROM ACTINOBACILLUS PLEUROPNEUMONIAE
; FILE REFERENCE: PC9854A
; CURRENT APPLICATION NUMBER: US/10/770,824
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: US/09/418,980
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1922
; TYPE: DNA
; ORGANISM: Actinobacillus pleuropneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (614)..(1705)
US-10-770-824-7

Query Match   15.4%;   Score 159.4;   DB 18;   Length 1922;
Best Local Similarity 56.2%;   Pred. No. 5.1e-41;
Matches 339;   Conservative 0;   Mismatches 261;   Indels 3;   Gaps 2;

```
Qy   414  TTCCCCAGTATTTGCTGGCGGGGTAGAGTGGGCTTACTGTGACATCGTTACTCGTCTACCCGTCT  473
          ||| ||  | || ||| |||  || ||||  || ||  || || ||  ||||
Db  1096  TTCTTTAATTTTAGGTGCGGGGTGTTGAGTGTTGAGTACGGCGCACGTGT  1155

Qy   474  GGAATACCAGTGGGTTAACACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAA  533
          |||||||  |||| |||| |||  |||| |  ||| | |  ||  ||| ||| | ||||
Db  1156  TGAATACCAATGGTTAAACAACCACGGCAGGTAAAGCAAGCAAGCTACTCTACTTTAAATCGTATGGG  1215

Qy   534  CGGCA-TGCTGAGCCTGGCGGTTTCCTACCGGTTCGGTCAGGAAGATGCTGC--ACCGGT  590
          ||||  ||||| | |||| | |  |  |||||| ||||  ||| |  || ||    ||||
Db  1216  TGCAACTGACTACCGTTCGGATATCAGTTCGGTTTAAGCTACCGTTTCGG  1275

Qy   591  TGTTGCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTC  650
          |  ||  | |||| ||| | ||| || |  | ||  ||||  |||| ||| |||
Db  1276  TCAAGGTGCGGTACCGGTTGCAGCTGCAGTTGAAACTAAAAACTTCGCATTCAGCTC  1335

Qy   651  TGACGTTCGTGTTCAACTTCAACAAAGGTACCCTGAAGCTACCTGAGA  710
          ||| |||  |||| |||| ||  |||||  ||  || ||  |||| ||
Db  1336  TGACGTATTATTCGCATTCGGTAAATCAAACTAAAAACTTCGGTTCTGGA  1395

Qy   711  TCAGCTGTACACTCAGCTGAGCAACAATGATCCGAAAGACGGTTCCGCTGTTGTTCTGGG  770
          ||||| ||| || |||| ||||  |||| ||  | |||| |||| |||||  || ||||
Db  1396  TGCAATGCAAACCGAAATCAATAACGCAGGTTTATCAAAGTCGCGATCCAAGTAAACGG  1455

Qy   771  CTACACCGACCGGTTCCGAAGCTTACACAACCAGCTCCGGAAGCTGCTCA  830
          |||||||| ||  | ||||| ||||| |||||||  |  ||||| |||| |
Db  1456  TTACACCGACCGTATCGGTAAAGAAGCTTCAAACTTAAAACTTTCACAACGTCGTGCGGA  1515

Qy   831  GTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTCGGCTCCGCGG  890
          || ||| || |||| |||  ||||| |||| ||| | ||  | | |||
Db  1516  AACAGTAGCTAACTAACTGTTTCTAAAGGTGCTCCGGCAGCAGTAACTGACAGTAGG  1575

Qy   891  CATGGGGTGAATCCAACCGGTTACTGCGACACACGTGTAACGGCGTCGGCTGC  950
          |||||  ||| ||||  | |||||| | || ||  || |||| ||| ||  ||
Db  1576  TTACGGTGAAGCAAACCCTGTAACCGGCGCCGCAACATGTGACAACGTAACTCGAGTAGG  1635
```

Qy   547 CTGGGCGTTTCCTACCGCTTCGGTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGA 606
             |||| | ||| ||||||||||||||||| || ||||||| ||||| || ||||||
Db   589 GGTTCTGTAACTGCTGCTGGTTTA---------TCTTACCGGTTTTGGTCAAAGTGCACCA 636

Qy   607 GCTCCCGCTCCGGAAGTGGCTACCAGCACTTCACCCTGAAGTCTCGTTGTTCGTTCAAC 666
             || || | || | ||| ||||| | | | | || |||||
Db   637 GTTGTTGAACCTAAGGTTGTTGTCAAAAACATTTGCATTAAATTCAGATGTTACTTTCGCA 696

Qy   667 TTCAACAAAGTACCCTGGAAACCGGAAGGTGCTCGGATCAGCTGTACACTCAG 726
             ||||||| | ||| | | |||| || | ||
Db   697 TTTGGTAAAGCAAATTTACGTCCAGAAGCACAAAAATGTATTTAGACGGTATTTATGGTGAA 756

Qy   727 CTGAGCAACATGGATCCGAAAGACAGGGTTCCGGTGTTCCGTGTTGTTCTCGGGCTACACCGACCGCATC 786
             | || || || ||| || | ||| | || ||||| ||||||
Db   757 ATCGCACAGTTAAAAATCAGTACAAGT------AGATGTTGCTGTTATACTGACCGTATT 810

Qy   787 GGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAAACGTGCTCAGTCCGTTGTTGACTAC 846
             || ||| |||| || |||| | | || ||| || ||| || |||| |||||||
Db   811 GGTAGCCGAAGCAGCCAACTTGAAATTATCACAACGTCGTGCTGATACTGTGGCTAACTAC 870

Qy   847 CTGGTTGCTAAAGGCCATCCCGCTGGCAAAATCTCCGCTCGCGGGCATGGGTGAATCCAAC 906
             | || |||| |||| | ||||| | || | | ||| || || ||| |||
Db   871 TTAGTTTCTAAAGGTGTTGCTCAAGAAGTGATTTCTTCAACAGGTTATGGTGAAGCCGAAC 930

Qy   907 CCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGGCTTGCCCTGATCGATTGCCTG 966
             | || | ||| || || ||| ||| || |||| || | || ||
Db   931 CCAGTAACTGGTCGGAAATGTGATACGGTTAAAGGTCGCAAGACAATAAATCCGCTTGTTTA 990

Qy   967 GCTCCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAA 1011
             || ||| |||| ||||| ||| ||| ||||| ||||| ||
Db   991 GCAGACGGATCGTCGTGTAGAGATTATTCCAGTTAAAGGTGAAGAGTAA 1035

Search completed: January 19, 2005, 18:13:24
Job time : 677 secs

Qy   951 CCTGATCGATTGCCTGGCTCCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGA 1010
             ||| ||| ||| ||| ||||| ||||| |||||| |||||||||||||
Db 1636 ATTAATCGCTGCTTGCTTAGCACCGGATCGTCGTGTTGAAGTTCAAGTGTTCAAGGTACTAAAGA 1695

Qy 1011 AGT 1013
             |||
Db 1696 AGT 1698

RESULT 15
US-10-336-840-20
; Sequence 20, Application US/10336840
; Publication No. US20030219454A1
; GENERAL INFORMATION:
; APPLICANT: TERRY, TAMSIN DEBORAH
; APPLICANT: TSENG, HSING-JU
; APPLICANT: HOBB, RHONDA IVY
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: DOWNES, JOHN
; TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
; FILE REFERENCE: 37955-0007
; CURRENT APPLICATION NUMBER: US/10/336,840
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: PCT/AU01/00822
; PRIOR FILING DATE: 2001-07-06
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 20
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Haemophilus paragallinarum
US-10-336-840-20

Query Match   13.0%; Score 135; DB 15; Length 1035;
Best Local Similarity 49.8%; Pred. No. 3.8e-33;
Matches 500; Conservative 0; Mismatches 475; Indels 30; Gaps 5;

Qy   16 GTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCC 75
             | | || | |||||| || ||| || |||||||| || |||| |||
Db   52 GTAGCACAAGCTGCACCACAGCAAACACTTTCTATGCTGTCTGTGCAAAAGCGGGGCTGGGCA 111

Qy   76 CAGTATCACGACACCGGTTTCTACGGTAACGGTTCCAGAACAACAACGGTTCCGACCCGT 135
               | | ||| |||| | || || | | ||| ||| | |||||||
Db   112 TCTTTCCACGATGGTTAAACCAATTTGAAAACTCACAAAAATGCAGATGGTACATTGCGT 171

Qy   136 AACGATCAGCTTGGTGCTGCGTTCGGTGCGTTCGGTGGTTTACCGGTTTACCTGGGTTTC 195
             || | | ||| || | | || |||| |||| | | ||| ||||| |||
Db   172 A--ATTCTGTAACTTATGGGTGTTCGGTGTTCGGTTACCAAATTACTGATAACTTCGCTGTT 228

Qy   196 GAAATGGGTTATGACTGGCGTGGCCGTGGGCCGTATGGCCATATAAAGGCAGCGTTGACAAGCGGTGCT 255
             ||| ||| || || ||| || | |||| | | | ||| | ||| ||| || ||
Db   229 GAGCTAGGTTATGATGAGTGCGGTGGCGTGTCAAGCGCGGTGCTAAGGCGGTGAAACTGTTATA 288

Qy   256 ---TTCAAAGCTCAGGGCGTTCAGCTACGGTGACCGCTAAACTGGGTTACCCGATCACTGACGAT 312
                |||||| || | | | | ||| | || || | |||| | | ||||||
Db   289 AAATACCACAAATCACGGAGCTTACTTAAGCTTAAAAGCAAGTTATCCAGTGCTTGAAGGA 348

Qy   313 CTGGACATCTCACACCCGTCTGGGCCGCGGCCATGGTTGGCGGCGCCTGACTCCAAAGGCAACTAC 372
             |||| | | |||| | | | | ||| | ||||| ||| | |||
Db   349 TTAGATGTTTATGCTGCGTTGCTCGCGTTGAGCAGCGTTGATTCGTTCTGATTATAAACCACCTAAA 408

Qy   373 GCTTCTACCGGCGTTCCGTTCACTGGCGTTCGTGCCGGAATACTGGTCCCAGTATTTGCTGGC 432
              || | || | |||| || | ||| | ||| || || || |||| | ||||| |||
Db   409 AGAGCAGCTCCTAATCAGACGCACGAACACATAGCTTAAAAGTTTCTCCAGTATAAACCACTAAA 468

Qy   433 GGCGTAGAGTGGGCTGTTAC------TCGTGACATCGCTACCCGTCTGGAATACCAGTGG 486
             || || || | | | | | | |||| || ||||| |||||
Db   469 GGTTTAGAGTATAACTTACCATCATTACCAGAACTTAAAGTTCTCCAGTATCAATGG 528

Qy   487 GTTAACAACATCGGCGACGCGGGCCACTGTGCCGTAGAGATCGAAGTTAAAGGCATGCTGAGC 546
             || | | | || | | || | | | |||||| || || | |||||||| | |||
Db   529 GTAAATAAAGTAGGAGCCGTGTGAAAAAGAGATGGTAGCCGGTGTTATACCAAGCATC 588

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 13:35:28 ; Search time 116 Seconds
(without alignments)
6341.953 Million cell updates/sec

Title: US-09-913-772-1
Perfect score: 1035
Sequence: 1 atgaaagcaatttcgtact.........taactcagccgcgggttaa 1035

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match %	Match Length	DB	ID	Description
1		1009.8	97.6	1158	4	US-09-489-039A-280	Sequence 280, App
2	c	1009.8	97.6	1251	4	US-09-489-039A-367	Sequence 367, App
3		1008	97.4	1008	3	US-08-721-979A-13	Sequence 13, Appl
4		1008	97.4	1008	4	US-09-654-289-13	Sequence 13, Appl
5		1008	97.4	1008	4	US-09-582-876-13	Sequence 13, Appl
6		1008	97.4	1008	4	US-10-091-257-13	Sequence 13, Appl
7		1007	97.3	1007	3	US-08-836-500A-1	Sequence 1, Appli
8		1007	97.3	1007	4	US-09-679-750-1	Sequence 1, Appli
9		724.4	70.0	730	2	US-08-743-637B-11	Sequence 11, Appl
10		724.4	70.0	730	3	US-08-526-840B-11	Sequence 11, Appl
11		537	51.9	537	3	US-08-836-500A-3	Sequence 3, Appli
12		537	51.9	537	4	US-09-679-750-3	Sequence 3, Appli
13		460.4	44.5	1155	4	US-09-543-681A-3750	Sequence 3750, Ap
14		224.8	21.7	1273	1	US-07-794-731A-1	Sequence 2, Appli
15		224.8	21.7	1273	1	US-07-794-731A-2	Sequence 2, Appli
16		216	20.9	216	4	US-08-836-500A-5	Sequence 5, Appli
17		216	20.9	216	4	US-09-679-750-5	Sequence 5, Appli
18		161	15.6	1095	4	US-09-809-665A-150	Sequence 150, App
19		159.4	15.4	1922	4	US-09-418-980-7	Sequence 7, Appli
20		159	15.4	159	3	US-08-836-500A-7	Sequence 7, Appli
21		159	15.4	159	4	US-09-679-750-7	Sequence 7, Appli
22		133.2	12.9	1720	1	US-08-457-997B-1	Sequence 1, Appli
23		133.2	12.9	1720	3	US-08-467-722A-1	Sequence 1, Appli
24		133.2	12.9	1720	4	US-09-451-184-1	Sequence 1, Appli
25		130	12.6	1830121	4	US-09-557-884-1	Sequence 1, Appli
26		130	12.6	1830121	4	US-09-643-990A-1	Sequence 1, Appli
27		130	12.6	1830121	4	US-10-329-960-1	Sequence 1, Appli
28		129.8	12.5	1319	4	US-09-418-980-9	Sequence 9, Appli
29		127.2	12.3	1059	3	US-08-476-102A-3	Sequence 3, Appli
30		125	12.1	1110	4	US-09-809-665A-152	Sequence 152, App
31		110	10.6	184	1	US-08-375-241-12	Sequence 12, Appl
32		110	10.6	184	5	PCT-US92-06617A-12	Sequence 12, Appl
33		92.4	8.9	169	1	US-08-375-241-13	Sequence 13, Appl
34		92.4	8.9	169	5	PCT-US92-06617A-13	Sequence 13, Appl
35	c	69.4	6.7	640681	4	US-09-790-988-1	Sequence 1, Appli
36		57.6	5.6	708	4	US-09-252-991A-3024	Sequence 3024, Ap
37	c	57.6	5.6	711	4	US-09-252-991A-2929	Sequence 2929, Ap
38		56.6	5.5	486	2	US-08-572-447C-10	Sequence 10, Appl
39		56.6	5.5	486	3	US-09-267-747-10	Sequence 10, Appl
40		56.6	5.5	645	2	US-08-572-447C-12	Sequence 12, Appl
41		56.6	5.5	645	3	US-09-267-747-12	Sequence 12, Appl
42		56.6	5.5	681	2	US-08-572-447C-14	Sequence 14, Appl
43		56.6	5.5	681	3	US-09-267-747-14	Sequence 14, Appl
44		56	5.4	534	4	US-09-252-991A-12949	Sequence 12949, A
45		56	5.4	819	4	US-09-252-991A-13110	Sequence 13110, A

ALIGNMENTS

RESULT 1
US-09-489-039A-280
; Sequence 280, Application US/0948039A
; Patent No. 661083
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 280
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-280

Query Match 97.6%;   Score 1009.8; DB 4;   Length 1158;
Best Local Similarity 99.3%;   Pred. No. 5.7e-277;
Matches 1014;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

```
Qy   15 CGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTC   74
          |||| |||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  138 CGTAGCGCAGGCCGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTC  197

Qy   75 CCAGTATCACGACACACCGGTTCTTACGGTTAACGGTTTCCAGAACAACAACGGTCCGACCCG  134
          ||||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db  198 CCAGTATCACGACGACCGGTTCTTACGGTTAACGGTTTCCAGAACAACAACAACGGTCCGACCCG  257

Qy  135 TAACGATCAGCTTGGTGCTGGTGGTCCGTTCGGTGGTGTTCCGTTGGTGTTTACCCGTACCTTCGGTTT  194
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  258 TAACGATCAGCTTGGTGCTGGTGGTCCGTTCAGTGGTGTTCCGTTGGTGTTTACCCGTACCTTCGGTTT  317

Qy  195 CGAAATGGGTTATGACTGGGCCTGGGCCGGCATATAAAGGCAGCGTTGACAACGGTGC  254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  318 CGAAATGGGTTATGACTGGGCCTGGGCCGGCATATAAAGGCAGCGTTGACAACGGTGC  377

Qy  255 TTTCAAAGCTCAGGGCGTTCAGCTGACCGGCTAAACTGGGTTACCGGTTACCGACACTGACGATCT  314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  378 TTTCAAAGCTCAGGGCGTTCAGCTGACCGGCTAAACTGGGTTACCGGTTACCGACACTGACGATCT  437

Qy  315 GGACATCTACACCCGTCTGGGCGGCCATGGTTTGGCGGCTAAACTGGTTTGGCGGCTCCAAAGGCAACTACGC  374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  438 GGACATCTACACCCGTCTGGGCGGCCATGGTTTGGCGGCTAAACTGGTTTGGCGGCTCCAAAGGCAACTACGC  497

Qy  375 TTCTACCGGCGTTTCCCGTAGCGAACGACACTGGCGGTTCCCCAGTATTGCTGGCGG  434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   498 TTCTACCGGCGTTCCCGTAGCGAACGACACTGGCGTTCCCCAGTATTTGCTGGCGG 557
Qy   435 CGTAGAGTGGGCTGTTACTCGTGACAATCGTCGTACCCGTTCTGTTGAGCGTTAACAA 494
Db   558 CGTAGAGTGGGCTGTTACTCGTGACATCGCTGCTACCCGTTCTGTTGAGCGTTAACAA 617
Qy   495 CATCCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATGCGGTTGGGCGT 554
Db   618 CATCGGCGACGCGGGCACCCGTGGGTACCCGTCCTGATAACGGCATGCGAGCCTGGGCGT 677
Qy   555 TTCCTACCGGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCGGCTCCGGCTCCGGC 614
Db   678 TTCCTACCGGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCGGCTCCGGCTCCGGC 737
Qy   615 TCCGGAAGTGGCTACCAAGACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAA 674
Db   738 TCCGGAAGTGGCTACCAAGACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAA 797
Qy   675 AGCTACCCTGAAAACCGGAAGTGCAGCAGCTCTGATCAGCTGTACACTCAGCTGAGCAA 734
Db   798 AGCTACCCTGAAAACCGGAAGTGCAGCAGCTCTGATCAGCTGTACACTCAGCTGAGCAA 857
Qy   735 CATGGATCCGAAAGACGGTTCCGCTGTGTTCTGGGCTACACCGACCGGCATGCGTTCCGA 794
Db   858 CATGGATCCGAAAGACGGTTCCGCTGTGTTCTGGGCTACACCGACCGGCATGCGTTCCGA 917
Qy   795 AGCTTACAACCAGCAGCTGTCTGAGAAATCGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 854
Db   918 AGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 977
Qy   855 TAAAGGCATCCCGGCTGGCAAACTCCGCTGCCGCGTGAAAGCTCGCGGCTGCTGGCTCCGGA 914
Db   978 TAAAGGCATCCCGGCTGGCAAAATCCGCTGCCGCGTGAAACCCGGTTAC 1037
Qy   915 TGGCAACACCTGTGACAACGTGAAGTAAAGGCTACAAAGAAGTTGTAACTGAAGCGGGGTTA 974
Db 1038 TGGCAACACCTGTGACAACGTGAAGTCGGGCTGTAAAGGCTACAAAGAAGTTGTAACTGAAGCGGGGTTA 1097
Qy   975 TCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGGCGGGGTTA 1034
Db 1098 TCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGGCGGGGTTA 1157

Qy 1035 A 1035
         |
Db 1158 A 1158

RESULT 2
US-09-489-039A-367/c
; Sequence 367, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 367
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-367

Query Match   97.6%; Score 1009.8; DB 4; Length 1251;
Best Local Similarity 99.3%; Pred. No. 5.9e-277;
Matches 1014; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy   15 CGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTC 74

(second column)

Db 1165 CGTACGCAGGCCGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTC 1106
Qy   75 CCAGTATCACGACACCGGTTCTACGGTGCGTTCCAGGTTAACCGGTTCCGACCCG 134
Db 1105 CCAGTATCACGACACCGGTTCTACGGTAACGGTTCCAGGAACAACAACAACGGTTCCGACCCG 1046
Qy   135 TAACGATCAGCTTGGTGTCGTGGTTCGGTGCGTTCGGTGGTTAACCCGTTACCTCGGTT 194
Db 1045 TAACGATCAGCTTGGTGTCGTGGTGCGTTCGGTGGTTAACCCGTTACCTCGGTT 986
Qy   195 CGAAATGGTTATGACTGGCTGGCCGTATGGCCATTATAAAGGCAGCGTTGACAACGGTGC 254
Db   985 CGAAATGGGTTATGACTGGCCGTATGGCCATTATAAAGGCAGCGTTGACAACGGTGC 926
Qy   255 TTTCAAAGCTCAGGGCGTTCAGCTGACCGTTCGACGTTCTGTTGTTCAACTTCAACTA 314
Db   925 TTTCAAAGCTCAGGGCGTTCAGCTGACCGTTAAACTGGGTTACCCGGTTCCGACGATCT 866
Qy   315 GGACATCTACACCCGTCGAAGGCTCAGCAGGCTCTGATCAGCTGTACACTCAGCTGAGCAA 374
Db   865 GGACATCTACACCCGTGGGCGCGCTGACTCAGCAGGCAACTACGC 806
Qy   375 TTCTACCGGCGGTTCCGCGTTTCGGTCGGTGCTACACCGACCGGCATGCGGCGG 434
Db   805 TTCTACCGGCGGTTTCCGCGTTTCGGTCGGTGCTACACCGACCGGCATGCGGCGG 746
Qy   435 CGTAGAGTGGGCTGTTACTCGTGACATGCTGTTACCCGTTCTGTTGAGCGTTAACAA 494
Db   745 CGTAGAGTGGGCTGTTACTCGTGACATCGCTGCTACCCGTTCTGTTGAGCGTTAACAA 686
Qy   495 CATCCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATGCGGGTTGGGCGT 554
Db   685 CATCGGCGACGCGGGCACCCGTGGGTACCCGTCCTGATAACGGCATGCGGGTTGGGCGT 626
Qy   555 TTCCTACCGGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCGGCTCCGGC 614
Db   625 TTCCTACCGGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCGGCTCCGGC 566
Qy   615 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAA 674
Db   565 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAA 506
Qy   675 AGCTACCCTGAAAACCGGAAGTGCTGACGAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAA 734
Db   505 AGCTACCCTGAAAACCGGAAGTGCTGACGAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAA 446
Qy   735 CATGGATCCGAAAAGACACGGTTCCGCTGTTCGGTGCTGTGTTCTGGGCTACACCGCATCGGTTCCGA 794
Db   445 CATGGATCCGAAAGACACGGTTCCGCTGTTCGGTGCTGTGTTCTGGGCTACACCGCATCGGTTCCGA 386
Qy   795 AGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 854
Db   385 AGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 326
Qy   855 TAAAGGCATCCCGGCTGGCAAAATCTCCGCTGCCGCGGTGAATCCAACCCGGTTAC 914
Db   325 TAAAGGCATCCCGGCTGGCAAAATCTCCGCTGCCGCGGTGAATCCAACCCGGTTAC 266
Qy   915 TGGCAACACCTGTGACAACGTGAAAGCTGCCCTGATCGATTGCCTGGCTCCGGA 974
Db   265 TGGCAACACCTGTGACAACGTGAAAGCTGCCCTGATCGATTGCCTGGCTCCGGA 206
Qy   975 TCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGGCGGGGTTA 1034
Db   205 TCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGGCGGGGTTA 146

Qy 1035 A 1035
         |
Db   145 A 145

RESULT 3
US-08-721-979A-13

```
Db  241  GGGGTTCAGCTGCGACCGCTAAACTGGGTTGGCGGCTGACATCACTGACGATCTGACGATCTACACC  300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  328  CGTCTGGGCGGCCCGCATGGTTGGCGGCTGACTCCAAAGGCAACTACGCTTCACCGGCGTT  387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CGTCTGGGCGGCCCGCATGGTTGGCGGCTGACTCCAAAGGCAACTACGCTTCACCGGCGTT  360
```

```
Qy  388  TCCCGTAGGGAACACGACACTGGGGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  447
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  TCCCGTAGGGAACACGACACTGGGGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  420
```

```
Qy  448  GTTACTCGTGACATGCTACCCGTCTGGAATACACCAGTGGTTAACAACATCGGCGACGCG  507
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GTTACTCGTGACATGCTACCCGTCTGGAATACACCAGTGGTTAACAACATCGGCGACGCG  480
```

```
Qy  508  GGCACTGTGGGTACCCGTCCTGATAACGGCGTTCGAGCCTGGGGGTTTCCTACCGCTTC  567
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  GGCACTGTGGGTACCCGTCCTGATAACGGCGTTCGAGCCTGGGGGTTTCCTACCGCTTC  540
```

```
Qy  568  GGTCAGGAAGATGTCCACCGGTTTGTGTCCGGCTCCGGCTCCGGAAGTGGCT  627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GGTCAGGAAGATGTCCACCGGTTTGTGTCCGGCTCCGGCTCCGGAAGTGGCT  600
```

```
Qy  628  ACCAAGCACTCACCCTGAAGTCGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA  687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  ACCAAGCACTCACCCTGAAGTCGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA  660
```

```
Qy  688  CCGGAAGGTCAGCAGCCTCGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA  747
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  CCGGAAGGTCAGCAGCCTCGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA  720
```

```
Qy  748  GACGGTTCCGCTGTTGTTCGGGTTCGGGGTACACCGACCGCATCGGTTCCGAAGTTACAACCAG  807
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GACGGTTCCGCTGTTGTTCGGGTTCGGGGTACACCGACCGCATCGGTTCCGAAGTTACAACCAG  780
```

```
Qy  808  CAGCTGTCTGAGAAACGTGCTCAGTCCGGTTGTTGACTACTCGGTTGCTAAAGGCATCCCCG  867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  CAGCTGTCTGAGAAACGTGCTCAGTCCGGTTGTTGACTACTCGGTTGCTAAAGGCATCCCCG  840
```

```
Qy  868  GCTGCAAAATCTCCGCTCGGCGCATGGGTGAATCCAACCGGTTACTGGCAACACCTGT  927
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  GCTGCAAAATCTCCGCTCGGCGCATGGGTGAATCCAACCGGTTACTGGCAACACCTGT  900
```

```
Qy  928  GACAACGTGAAAGCTGCGGCTGCCCGATCGATTGCCTGCCTGGATCGTCGTGTAGAG  987
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  GACAACGTGAAAGCTGCGGCTGCCCGATCGATTGCCTGCCTGGATCGTCGTGTAGAG  960
```

```
Qy  988  ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGTTAA  1035
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGTTAA  1008
```

RESULT 4
US-09-654-289-13

; Sequence 13, Application US/0954289
; Patent No. 6410030
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   APPLICANT: N'Guyen, Ngoc Thien
;   APPLICANT: Baussant, Thierry
;   APPLICANT: Trudel, Michel
;   TITLE OF INVENTION: PEPTIDE FRAGMENT OF RESPIRATORY
;   TITLE OF INVENTION: SYNCYTIAL VIRUS PROTEIN G, IMMUNOGENIC AGENT, PHARMACEUTICAL
;   TITLE OF INVENTION: COMPOSITION CONTAINING IT AND PREPARATION PROCESS
;   NUMBER OF SEQUENCES: 75
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Gordon W. Hueschen
;   STREET: 715 The "H" Bldg., 310 East Michigan
;   STREET: Avenue
;   CITY: Kalamazoo
;   STATE: MI
;   COUNTRY: USA
;   ZIP: 49007

; Sequence 13, Application US/08721979A
; Patent No. 6113911
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   APPLICANT: N'Guyen, Ngoc Thien
;   APPLICANT: Baussant, Thierry
;   APPLICANT: Trudel, Michel
;   TITLE OF INVENTION: PEPTIDE FRAGMENT OF RESPIRATORY
;   TITLE OF INVENTION: SYNCYTIAL VIRUS PROTEIN G, IMMUNOGENIC AGENT, PHARMACEUTICAL
;   TITLE OF INVENTION: COMPOSITION CONTAINING IT AND PREPARATION PROCESS
;   NUMBER OF SEQUENCES: 75
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Gordon W. Hueschen
;   STREET: 715 The "H" Bldg., 310 East Michigan
;   STREET: Avenue
;   CITY: Kalamazoo
;   STATE: MI
;   COUNTRY: USA
;   ZIP: 49007
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   SOFTWARE:
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/721,979A
;   FILING DATE: October 4, 1996
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 94 04009
;   FILING DATE: 06-APR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Hueschen, Gordon W.
;   REGISTRATION NUMBER: 16,157
;   REFERENCE/DOCKET NUMBER: PF57PCTUS/dln
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 616-382-0030
;   TELEFAX: 616-382-2030
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1008 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..1008
;   OTHER INFORMATION: /note= "name : P40"
US-08-721-979A-13

Query Match   97.4%; Score 1008; DB 3; Length 1008;
Best Local Similarity 100.0%; Pred. No. 1.8e-276;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   28  GCTCCGAAAGATAACAACACCTGGGTATGCGACGTGGTAAACTGGGTTGGGTTGGTCTCCCAGTATCACGAC  87
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GCTCCGAAAGATAACAACACCTGGGTATGCGACGTGGTAAACTGGGTTGGTCTCCCAGTATCACGAC  60
```

```
Qy   88  ACCGGTTTCTACGGTAACGGTTCCAGACAACAACACCGTCCGACCCGGTAACGATCAGCTT  147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  ACCGGTTTCTACGGTAACGGTTCCAGAACAACAACAACCGTCCGACCCGGTAACGATCAGCTT  120
```

```
Qy  148  GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGGTTACCTCGGTTTCGAAATGGGTTAT  207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGGTTACCTCGGTTTCGAAATGGGTTAT  180
```

```
Qy  208  GACTGGCTGGGCCGTATGGCGATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG  267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GACTGGCTGGGCCGTATGGCGATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG  240
```

```
Qy  268  GGCGTTCAGCTGCGACCGCTAAACTGGGTTACCCGATCACTGACGATCTGACATCTCACACC  327
```

```
Qy   568 GGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGGCTCCGGCTCCGGAAGTGCT 627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGGCTCCGGCTCCGGAAGTGCT 600

Qy   628 ACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCCTGAAA 687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCCTGAAA 660

Qy   688 CCGGAAGGTCAGCAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 747
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCGGAAGGTCAGCAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 720

Qy   748 GACGGTTCCGCTGTTGTTCTGGGCTACACCGACCGGTTCCGAAGCTTACAACCAG 807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGTTCCGCTGTTGTTCTGGGCTACACCGACCGGTTCCGAAGCTTACAACCAG 780

Qy   808 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 840

Qy   868 GCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCAACCAACCGTTACTGGCAACACTGT 927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCAACCAACCGTTACTGGCAACACTGT 900

Qy   928 GACAACGTGAAAGCTCGCGCCTGCCTGATCGATTGCCTGGCTCCGGATCGTCGTGTAGAG 987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GACAACGTGAAAGCTCGCGCCTGCCTGATCGATTGCCTGGCTCCGGATCGTCGTGTAGAG 960

Qy   988 ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGGTTAA 1035
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGGTTAA 1008
```

RESULT 5
US-09-582-876-13
; Sequence 13, Application US/09582876
; Patent No. 6537556
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   N'Guyen, Ngoc Thien
;   Baussant, Thierry
;   Trudel, Michel
;   TITLE OF INVENTION: PEPTIDE FRAGMENT OF RESPIRATORY
;   SYNCYTIAL VIRUS PROTEIN G, IMMUNOGENIC AGENT, PHARMACEUTICAL
;   COMPOSITION CONTAINING IT AND PREPARATION PROCESS
;   NUMBER OF SEQUENCES: 75
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Gordon W. Hueschen
;   STREET: 715 The "H" Bldg., 310 East Michigan
;   Avenue
;   CITY: Kalamazoo
;   STATE: MI
;   COUNTRY: USA
;   ZIP: 49007
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   (EPO)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/582,876
;   FILING DATE: 30-Jun-2000
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/721,979
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Hueschen, Gordon W.
;   REGISTRATION NUMBER: 16,157
;   REFERENCE/DOCKET NUMBER: PF57PCTUS/dln
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 616-382-0030

;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/654,289
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/721,979
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Hueschen, Gordon W.
;   REGISTRATION NUMBER: 16,157
;   REFERENCE/DOCKET NUMBER: PF57PCTUS/dln
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 616-382-0030
;   TELEFAX: 616-382-2030
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1008 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1...1008
;   OTHER INFORMATION: /note= "name : P40"
US-09-654-289-13

Query Match   97.4%;   Score 1008;   DB 4;   Length 1008;
Best Local Similarity   100.0%;   Pred. No. 1.8e-276;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    28 GCTCCGAAAGATAACACCCTGGTATGCGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC 87
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GCTCCGAAAGATAACACCCTGGTATGCGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC 60

Qy    88 ACCGGTTTCTACGGTAACGGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT 147
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ACCGGTTTCTACGGTAACGGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT 120

Qy   148 GGTGCTGGTGGCGTTCGGTGGTTACCAGGTTAACCCGTTCGGTTTCGAAATGGGTTAT 207
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGTGCTGGTGGCGTTCGGTGGTTACCAGGTTAACCCGTTCGGTTTCGAAATGGGTTAT 180

Qy   208 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG 267
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG 240

Qy   268 GGGTTCAGCTGGTTACCCGGTACACCGCTAAACTGGGTTACCCGTTACCGGACATCTACACC 327
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGGTTCAGCTGGTTACCCGGTACACCGCTAAACTGGGTTACCCGTTACCGGACATCTACACC 300

Qy   328 CGTCTGGGCGGCATGTTTGCGCGCGGACTCCAAAGGCAACTACGCTTCTACCGGCGTT 387
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGTCTGGGCGGCATGTTTGCGCGCGGACTCCAAAGGCAACTACGCTTCTACCGGCGTT 360

Qy   388 TCCCGTAGCGAACAACGACACTGGCGTTTCCCAGTATTTGCTGGCGGGCTAGAGTGGGCT 447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCCCGTAGCGAACAACGACACTGGCGTTTCCCAGTATTTGCTGGCGGGCTAGAGTGGGCT 420

Qy   448 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 507
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 480

Qy   508 GGCACTGTGGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGTTCCTACCGGCTTC 567
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GGCACTGTGGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGTTCCTACCGGCTTC 540
```

```
;   TELEFAX: 616-382-2030
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1008 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1008
;     OTHER INFORMATION: /note= "name : P40"
;   SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-582-876-13

Query Match        97.4%;  Score 1008;  DB 4;  Length 1008;
Best Local Similarity  100.0%;  Pred. No. 1.8e-276;
Matches 1008;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      28 GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   60

Qy      88 ACCGGTTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT  147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ACCGGTTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT  120

Qy     148 GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTACCTCCGTTTCGAAATGGGTTAT  207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTACCTCCGTTTCGAAATGGGTTAT  180

Qy     208 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG  267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCTTTCAAAGCTCAG  240

Qy     268 GGGGTTCAGCTGACCGGCTAAACTGGGTTACCCGGTTACCCGACATCTGACGACATCTACACC  327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGGGTTCAGCTGACCGGCTAAACTGGGTTACCCGGTTACCCGACATCTGACGACATCTACACC  300

Qy     328 CGTCTGGGCGGCCATGGTTGGCGCGTGACTGCTCCAAAGGCAACTACGCTTCTACCGGCGTT  387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGTCTGGGCGGCCATGGTTGGCGCGTGACTGCTCCAAAGGCAACTACGCTTCTACCGGCGTT  360

Qy     388 TCCCGTAGCGAACAACGACACTGCGCGTTTCCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TCCCGTAGCGAACAACGACACTGCGCGTTTCCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  420

Qy     448 GTTACTCGTGACATCGCTACCCGTCTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG  507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTTACTCGTGACATCGCTACCCGTCTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG  480

Qy     508 GGCACTGTGGGTACCCGTCCTGATAACCGGCATGCTGAGCCTGGCGTTCCTACCGCTTC  567
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GGCACTGTGGGTACCCGTCCTGATAACCGGCATGCTGAGCCTGGCGTTCCTACCGCTTC  540

Qy     568 GGTCAGGAAGATGCTGACCACGGTTGTTGCTCCGGCTCCGGAAGTGCT  627
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGTCAGGAAGATCCTGACACCGGTTGTTGCTCCGGCTCCGGAAGTGCT  600

Qy     628 ACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA  687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGCTACCCTGAAA  660

Qy     688 CCGGAAGGTCAGCAGGCTCTGGTACACTGTACACCGTGACTGAGCAACATGGATCCGAAA  747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCGGAAGGTCAGCAGGCTCTGGTACACTGTACACCGTGACTGAGCAACATGGATCCGAAA  720

Qy     748 GACGGGTTCCGCTGTGTTCTGGGCTACACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GACGGGTTCCGCTGTGTTCTGGGCTACACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  780

Qy     808 CAGCTGCTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG  867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     781 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG  840

Qy     868 GCTGCAAAATCTCCGCTCGCGGCATGGGCATGGGCGTGAATCCAACCGGTTACTGCAACACCTGT  927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GCTGCAAAATCTCCGCTCGCGGCATGGGCATGGGCGTGAATCCAACCGGTTACTGCAACACCTGT  900

Qy     928 GACAACGTGAAAGCTGCGCGGCTGCCCTGACATCGATTGCCTGATCGTCCGTGTAGAG  987
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GACAACGTGAAAGCTGCGCGGCTGCCCGATCGATTGCCTGATCGTCCGTGTAGAG  960

Qy     988 ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGTTAA  1035
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGGGTTAA  1008
```

```
RESULT 6
US-10-091-257-13
; Sequence 13, Application US/10091257
; Patent No. 6616930
;   GENERAL INFORMATION:
;     APPLICANT: Binz, Hans
;                N'Guyen, Ngoc Thien
;                Baussant, Thierry
;                Trudel, Michel
;   TITLE OF INVENTION: PEPTIDE FRAGMENT OF RESPIRATORY
;                SYNCYTIAL VIRUS PROTEIN G, IMMUNOGENIC AGENT, PHARMACEUTICAL
;                COMPOSITION CONTAINING IT AND PREPARATION PROCESS
;
;   NUMBER OF SEQUENCES: 75
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Gordon W. Hueschen
;     STREET: 715 The "H" Bldg., 310 East Michigan
;                Avenue
;     CITY: Kalamazoo
;     STATE: MI
;     COUNTRY: USA
;     ZIP: 49007
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;                (EPO)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/091,257
;     FILING DATE: 05-Mar-2002
;     CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/08/721,979A
;     FILING DATE: October 4, 1996
;     APPLICATION NUMBER: FR 94 04009
;     FILING DATE: 06-APR-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Hueschen, Gordon W.
;     REGISTRATION NUMBER: 16,157
;     REFERENCE/DOCKET NUMBER: PF57PCTUS/dln
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 616-382-0030
;     TELEFAX: 616-382-2030
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1008 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1008
;     OTHER INFORMATION: /note= "name : P40"
;   SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-091-257-13

Query Match        97.4%;  Score 1008;  DB 4;  Length 1008;
```

US-08-836-500A-1
; Sequence 1, Application US/0836500A
; Patent No. 6197929
; GENERAL INFORMATION:
; APPLICANT: Binz, Hans
; APPLICANT: Baussant, Thierry
; APPLICANT: Haeuw, Jean-Francois
; APPLICANT: Nguyen Ngoc, Thien
; TITLE OF INVENTION: Carrier Protein Having an Adjuvant
; TITLE OF INVENTION: Effect, Immunogenic Complex Containing It, Process for
; TITLE OF INVENTION: Their Preparation, Nucleotide Sequence and Vaccines
; Patent No. 6197929
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
; STREET: 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,500A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: PIE1514P0180US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1007
;
US-08-836-500A-1

Query Match 97.3%; Score 1007; DB 3; Length 1007;
Best Local Similarity 100.0%; Pred. No. 3.4e-276;
Matches 1007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   28 GCTCCGAAAGATAACACCTGGTATGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   87
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCTCCGAAAGATAACACCTGGTATGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   60

Qy   88 ACCGGTTTCTACGGTAACGGTTCCAGACAACAACACGGTCCGACCCGTAACGATCAGCTT 147
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACCGGTTTCTACGGTAACGGTTCCAGACAACAACACGGTCCGACCCGTAACGATCAGCTT 120

Qy 148 GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGTACCTCGGTTTCGAAATGGGTTAT 207
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGTACCTCGGTTTCGAAATGGGTTAT 180

Qy 208 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGGTGACAACGGTGCTTTCAAAGCTCAG 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGGTGACAACGGTGCTTTCAAAGCTCAG 240

Qy 268 GGCGTTCAGCTGACCCGCTAAACTGGGTTACCCGGTTGACACTCTGGACATCTACACC 327
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCGTTCAGCTGACCCGCTAAACTGGGTTACCCGGTTGACACTCTGGACATCTACACC 300

US-08-836-500A-1

Best Local Similarity 100.0%; Pred. No. 1.8e-276;
Matches 1008; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   28 GCTCCGAAAGATAACACCTGGTATGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   87
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCTCCGAAAGATAACACCTGGTATGGAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   60

Qy   88 ACCGGTTTCTACGGTAACGGTTCCAGACAACAACACGGTCCGACCCGTAACGATCAGCTT 147
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ACCGGTTTCTACGGTAACGGTTCCAGACAACAACACGGTCCGACCCGTAACGATCAGCTT 120

Qy 148 GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGTACCTCGGTTTCGAAATGGGTTAT 207
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GGTGCTGGTGCGTCGGTGGTTACCAGGTTAACCCGTACCTCGGTTTCGAAATGGGTTAT 180

Qy 208 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGGTGACAACGGTGCTTTCAAAGCTCAG 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGGTGACAACGGTGCTTTCAAAGCTCAG 240

Qy 268 GGCGTTCAGCTGACCCGCTAAACTGGGTTACCCGGTTGACACATCTGGACATCTACACC 327
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 GGCGTTCAGCTGACCCGCTAAACTGGGTTACCCGGTTGACACATCTGGACATCTACACC 300

Qy 328 CGTCTGGGCCGGCATGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTTCTACCGGCGTT 387
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CGTCTGGGCCGGCATGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTTCTACCGGCGTT 360

Qy 388 TCCCGTAGCGAACACGACACTGGCGTTTCCCCAGTATTGCTGGGCGGCGTAGAGTGGGCT 447
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 TCCCGTAGCGAACACGACACTGGCGTTTCCCCAGTATTGCTGGGCGGCGTAGAGTGGGCT 420

Qy 448 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 507
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAACATCGGCGACGCG 480

Qy 508 GGCACTGTGGGTACCCGTCCTGATAACGGCATGCGTTCTGAGCCTGGGCGTTTCCTACCGCTTC 567
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GGCACTGTGGGTACCCGTCCTGATAACGGCATGCGTTCTGAGCCTGGGCGTTTCCTACCGCTTC 540

Qy 568 GGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGGCTCCGGCTCCGGAAGTGGGCT 627
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGGCTCCGGCTCCGGAAGTGGGCT 600

Qy 628 ACCAAGCACTTCACCCTGAAGTCTGACGTTCGTTCAACTTCAACAAAGTACCCTGAAA 687
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 ACCAAGCACTTCACCCTGAAGTCTGACGTTCGTTCAACTTCAACAAAGTACCCTGAAA 660

Qy 688 CCGGAAGGTCAGCAGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 747
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 CCGGAAGGTCAGCAGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 720

Qy 748 GACGGTTCCGCGTTGTTGTTCCGCGTGTTGTTGTTCCGGGCTTCACACCGACCGCATCCG 807
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GACGGTTCCGCGTTGTTGTTCCGCGTGTTGTTGTTCCGGGCTTCACACCGACCGCATCCG 780

Qy 808 CAGCTGCTCTGAGAAACGTGCTCAGTCGCTTCAGTCCGTTGTTGACTACCTTGTTCTAAAGGCATCCCG 867
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 CAGCTGCTCTGAGAAACGTGCTCAGTCGCTTCAGTCCGTTGTTGACTACCTTGTTCTAAAGGCATCCCG 840

Qy 868 GCTGGCAAATCTCCGCCGCCGCCGGTTACTGGGTCAACACCTGT 927
         ||||||||||||||||||||||||||||||||||||||||||||
Db 841 GCTGGCAAATCTCCGCCGCCGCCGGTTACTGGGTCAACACCTGT 900

Qy 928 GACAACGTGAAAGCTCGCGCGCTCCGGCTCCGGATCGTCGTGTAGAG 987
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 901 GACAACGTGAAAGCTCGCGCGCTCCGGCTCCGGATCGTCGTGTAGAG 960

Qy 988 ATCGAAGTTAAAGCTACAAAGGCTACAAAGAAGTTGTAACTCAGCCGGGCGGGGTTAA 1035
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 ATCGAAGTTAAAGCTACAAAGGCTACAAAGAAGTTGTAACTCAGCCGGGCGGGGTTAA 1008

RESULT 7

```
                                    COMPUTER: IBM PC compatible
                                    OPERATING SYSTEM: PC-DOS/MS-DOS
                                    SOFTWARE: PatentIn Release #1.0, Version #1.30
                                 CURRENT APPLICATION DATA:
                                    APPLICATION NUMBER: US/09/679,750
                                    FILING DATE: 08-Oct-2000
                                    CLASSIFICATION: <Unknown>
                                 PRIOR APPLICATION DATA:
                                    APPLICATION NUMBER: 08/836,500
                                    FILING DATE: <Unknown>
                                 ATTORNEY/AGENT INFORMATION:
                                    NAME: Katz, Martin L.
                                    REGISTRATION NUMBER: 25,011
                                    REFERENCE/DOCKET NUMBER: PIE1514P0180US
                                 TELECOMMUNICATION INFORMATION:
                                    TELEPHONE: 312-616-5400
                                    TELEFAX: 312-616-5460
                                 INFORMATION FOR SEQ ID NO: 1:
                                    SEQUENCE CHARACTERISTICS:
                                       LENGTH: 1007 base pairs
                                       TYPE: nucleic acid
                                       STRANDEDNESS: single
                                       TOPOLOGY: linear
                                    MOLECULE TYPE: cDNA
                                    FEATURE:
                                       NAME/KEY: CDS
                                       LOCATION: 1..1007
                                    SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-679-750-1

Query Match        97.3%;  Score 1007;  DB 4;  Length 1007;
Best Local Similarity   100.0%;  Pred. No. 3.4e-276;
Matches 1007;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy     28  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGTTGGTCCCAGTATCACGAC  87
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGTTGGTCCCAGTATCACGAC  60

Qy     88  ACCGGTTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCGGTAACGATCAGCTT  147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  ACCGGTTTCTACGGTAACGGTTTCCAGAACAACAACGGTCCGACCCGGTAACGATCAGCTT  120

Qy    148  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTCGGTTTCGAAATGGGTTAT  207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTCGGTTTCGAAATGGGTTAT  180

Qy    208  GACTGGCTGGGCCGTATGGCATCTAAAGGCAGCGTTGACAACGGTGCTTTCAAAGTCAG  267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  GACTGGCTGGGCCGTATGGCATCTAAAGGCAGCGTTGACAACGGTGCTTTCAAAGTCAG  240

Qy    268  GGCGTTCAGCTGACGACCGCTAAACTGGTTACCGGATCACTGACGCATCTACACC  327
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GGCGTTCAGCTGACGACCGCTAAACTGGTTACCGGATCACTGACGCATCTACACC  300

Qy    328  CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT  387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT  360

Qy    388  TCCCGTAGCGAACGACACTGGCGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCCCGTAGCGAACGACACTGGCGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  420

Qy    448  GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGTTAACAACATCGGCGACGCG  507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGTTAACAACATCGGCGACGCG  480

Qy    508  GGCACTGTGGGTACCCGTCCCTGATAACGGCATGCATGCTGAGCCTGGCGTTTCCTACCGCCTTC  567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGCACTGTGGGTACCCGTCCCTGATAACGGCATGCATGCTGAGCCTGGCGTTTCCTACCGCCTTC  540

Qy    568  GGTCAGGAAGATGCTCACCGGTTGTTGCTCCGGAAGTGGCT  627
           |||||||||||||||||||||||||||||||||||||||||
Db    541  GGTCAGGAAGATGCTCACCGGTTGTTGCTCCGGAAGTGGCT  600
```

```
Qy    328  CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT  387
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  CGTCTGGGCGGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGTT  360

Qy    388  TCCCGTAGCGAACGACACTGGCGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCCCGTAGCGAACGACACTGGCGTTCCCAGTATTTGCTGGCGGCGTAGAGTGGGCT  420

Qy    508  GGCACTGTGGGTACCCGTCCCTGATAACGGCATGCATGCTGAGCCTGGCGTTTCCTACCGCCTTC  567
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGCACTGTGGGTACCCGTCCCTGATAACGGCATGCATGCTGAGCCTGGCGTTTCCTACCGCCTTC  540

Qy    568  GGTCAGGAAGATGCTCACCGGTTGTTGCTCCGGAAGTGGCT  627
           |||||||||||||||||||||||||||||||||||||||||
Db    541  GGTCAGGAAGATGCTCACCGGTTGTTGCTCCGGAAGTGGCT  600

Qy    628  ACCAAGCACTTCACCCTGAAGTCTGTTCAACTTCAACAAAGCTACCCTGAAA  687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  ACCAAGCACTTCACCCTGAAGTCTGTTCAACTTCAACAAAGCTACCCTGAAA  660

Qy    688  CCGGAAGGTCAGCAGGCTCTGATCAGCTGTACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  747
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  CCGGAAGGTCAGCAGGCTCTGATCAGCTGTACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  720

Qy    748  GACGGTTCCGCTGTGTTGTTCTGGGCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  807
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GACGGTTCCGCTGTGTTGTTCTGGGCTACACCGACCGCATCGGTTCCGAAGCTTACAACCAG  780

Qy    808  CAGCTGTCTGAGAAACGTGCTCAGTCCGTTCAGTCCCGTTGTTGACTACCTGTTGCTAAAGGCATCCCG  867
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CAGCTGTCTGAGAAACGTGCTCAGTCCGTTCAGTCCCGTTGTTGACTACCTGTTGCTAAAGGCATCCCG  840

Qy    868  GCTGGCAAAATCTTCCGCTCGCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGT  927
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GCTGGCAAAATCTTCCGCTCGCGCGGCATGGGTGAATCCAACCCGGTTACTGGCAACACCTGT  900

Qy    928  GACAACGTGAAAGCTGAAAGCTCCGCGGCTGCCCTGATCGATTGCCTGGCTCCGGATCGTCGTGTAGAG  987
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  GACAACGTGAAAGCTGAAAGCTCCGCGGCTGCCCTGATCGATTGCCTGGCTCCGGATCGTCGTGTAGAG  960

Qy    988  ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGCGGTTA  1034
           ||||||||||||||||||||||||||||||||||||||||
Db    961  ATCGAAGTTAAAGGCTACAAAGAAGTTGTAACTCAGCCGCGGCGGTTA  1007
```

```
RESULT 8
US-09-679-750-1
; Sequence 1, Application US/09679750
; Patent No. 6780420
; GENERAL INFORMATION:
   APPLICANT: Binz, Hans
             Baussant, Thierry
             Haeuw, Jean-Francois
             Nguyen Ngoc, Thien
   TITLE OF INVENTION: Carrier Protein Having an Adjuvant
          Effect, Immunogenic Complex Containing It, Process for
          Their Preparation, Nucleotide Sequence and Vaccines
; Patent No. 6780420
   NUMBER OF SEQUENCES: 8
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
      STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
            4700
      CITY: Chicago
      STATE: Illinois
      COUNTRY: U.S.A.
      ZIP: 60601
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
```

```
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: DNA (genomic)
;      ORIGINAL SOURCE:
;      ORGANISM: Klebsiella pneumoniae
US-08-743-637B-11

Query Match      70.0%;  Score 724.4;  DB 2;  Length 730;
Best Local Similarity  99.9%;  Pred. No. 5.2e-196;
Matches   725;  Conservative   0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   276 GCTGACCGCTAAACTGGGTTACCGATCACTGACGATCTGACGATCTACACCCGTCTCGGG 335
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GCTGACCGCTAAACTGGGTTACCGATCACTGACGATCTGACGATCTACACCCGTCTCGGG 60

Qy   336 CGGCATGGTTTGGCGGCGCTGACTCCAAAGGCAACTACGCTTCTACCCGCGTTTCCCGTAG 395
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGGCATGGTTTGGCGGCGCTGACTCCAAAGGCAACTACGCTTCAACCGGCGTTTCCCGTAG 120

Qy   396 CGAACACGACACTGGCGTTCCCAGTATTTGCTGGCGGTAGAGTGGGCGTGTTACTCG 455
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGAACACGACACTGGCGTTCCCAGTATTTGCTGGCGGTAGAGTGGGCGTGTTACTCG 180

Qy   456 TGACATCGCTACCCGTCTGGAATAACCAGTGGGTTAACAACATCGGCGACGCGGGCACTGT 515
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGACATCGCTACCCGTCTGGAATAACCAGTGGGTTAACAACATCGGCGACGCGGGCACTGT 240

Qy   516 GGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGTTCCTACCCGCCTTCGGTCAGGA 575
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGTTCCTACCCGCCTTCGGTCAGGA 300

Qy   576 AGATGCTGCACCGGTTGTGCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCA 635
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGATGCTGCACCGGTTGTGCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCA 360

Qy   636 CTTCACCCTGAAGTCTGCAGTTCTGTTCAACTTCAACAAAGTACCCTGAAACCGGAAGG 695
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTTCACCCTGAAGTCTGCAGTTCTGTTCAACTTCAACAAAGTACCCTGAAACCGGAAGG 420

Qy   696 TCAGCAGGCTCTGGATCCGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAAGACGGTTC 755
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCAGCAGGCTCTGGATCCGATCAGCTGTACACTCAGCTGAGCAACAACATGGATCCGAAAGACGGTTC 480

Qy   756 CGCTGTTGTTCTGGGCTCACACCGACCGCATCGGTTCCGAAGTTACAACCAGCAGCTGTC 815
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CGCTGTTGTTCTGGGCTCACACCGACCGCATCGGTTCCGAAGTTACAACCAGCAGCTGTC 540

Qy   816 TGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAA 875
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAA 600

Qy   876 AATCTCCGGCTCCGGCGCATGGCAATCCAACCCGTTACTGGCAACACCTGTGACAACGT 935
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AATCTCCGGCTCCGGCGCATGGCAATCCAACCCGTTACTGGCAACACCTGTGACAACGT 660

Qy   936 GAAAGCTCGCGCTGCCCTGCGTCGATCGATTGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGT 995
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GAAAGCTCGCGCTGCCCTGCGTCGATCGATTGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGT 720

Qy   996 TAAAGG 1001
         ||||||
Db   721 TAAAGG 726
```

```
RESULT 10
US-08-526-840B-11
; Sequence 11, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
;      APPLICANT: BERGERON, Michel G.
;      APPLICANT: OUELLETTE, Marc
;      APPLICANT: ROY, Paul H.
```

```
Qy   628 ACCAAGCACTTCACCCTGAAGTCTGCAGTTCTGTTCAACTTCAACAAAGTACCCTGAAA 687
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACCAAGCACTTCACCCTGAAGTCTGCAGTTCTGTTCAACTTCAACAAAGTACCCTGAAA 660

Qy   688 CCGGAAGGTCAGCAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 747
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CCGGAAGGTCAGCAGGCTCTGGATCAGCTGTACACTCAGCTGAGCAACATGGATCCGAAA 720

Qy   748 GACGGTTCCGCTGTTGTTCTGGGCTCACACCGACCGCATCGGTTCCGAAGCTTACAACCAG 807
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGTTCCGCTGTTGTTCTGGGCTCACACCGACCGCATCGGTTCCGAAGCTTACAACCAG 780

Qy   808 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 867
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGCTAAAGGCATCCCG 840

Qy   868 GCTGGCAAAATCTCCGGCTCCGGCGCATGGGCAATCCAACCCGGTTACTGGCAACACCTGT 927
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GCTGGCAAAATCTCCGGCTCCGGCGCATGGGCAATCCAACCCGGTTACTGGCAACACCTGT 900

Qy   928 GACAACGTGAAAGCTCGCGCTGCCCTGCGTCGATCGATTGCCTGCTCCGGATCGTCGTAGAG 987
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GACAACGTGAAAGCTCGCGCTGCCCTGCGTCGATCGATTGCCTGCTCCGGATCGTCGTGTAGAG 960

Qy   988 ATCGAAGTTAAAGGCTACAAAGAAGAGTTGTAACTCAGCCGCGGCGGTTA 1034
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 ATCGAAGTTAAAGGCTACAAAGAAGAGTTGTAACTCAGCCGCGGCGGTTA 1007
```

```
RESULT 9
US-08-743-637B-11
; Sequence 11, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
;      APPLICANT: BERGERON, Michel G.
;      APPLICANT: PICARD, Francois J.
;      APPLICANT: OUELLETTE, Marc
;      APPLICANT: ROY, Paul H.
;      TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
;      TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
;      TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
;      TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS  ...
;      NUMBER OF SEQUENCES: 273
;      CORRESPONDENCE ADDRESS:
;      ADDRESSEE: QUARLES & BRADY
;      STREET: 411 EAST WISCONSIN AVENUE
;      CITY: MILWAUKEE
;      STATE: WISCONSIN
;      COUNTRY: USA
;      ZIP: 53202-4497
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/743,637B
;      FILING DATE: 04-NOV-1996
;      CLASSIFICATION: 435
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/526,840
;      FILING DATE: 11-SEP-1995
;      ATTORNEY/AGENT INFORMATION:
;      NAME: BAKER, Jean C.
;      REGISTRATION NUMBER: 35,433
;      REFERENCE/DOCKET NUMBER: 850586.90012
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (414) 277-5000
;      TELEFAX: (414)277-5591
; INFORMATION FOR SEQ ID NO: 11:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 730 base pairs
```

```
; TITLE OF INVENTION:    SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION:    AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION:    COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES
; TITLE OF INVENTION:    FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES:   177
; CORRESPONDENCE ADDRESS:
;     ADDRESSEE: QUARLES & BRADY
;     STREET:    411 East Wisconsin Avenue
;     CITY:      Milwaukee
;     STATE:     Wisconsin
;     COUNTRY:   USA
;     ZIP:       53202-4497
; COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER:    IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:    PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/526,840B
;     FILING DATE: 11-SEP-1995
;     CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/304,732
;     FILING DATE: 12-SEP-1994
; ATTORNEY/AGENT INFORMATION:
;     NAME: BAKER, Jean C.
;     REGISTRATION NUMBER: 35,433
;     REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (414) 277-5000
;     TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 11:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 730 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     ORIGINAL SOURCE:
;         ORGANISM: Klebsiella pneumoniae
US-08-526-840B-11

Query Match          70.0%;   Score 724.4;  DB 3;   Length 730;
Best Local Similarity 99.9%;   Pred. No. 5.2e-196;
Matches 725;   Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy   276 GCTGACCGCTAAACTGGTTACCGGATCACTGACGATCTGGACATCTACACCCGTCTGGG 335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GCTGACCGCTAAACTGGTTACCGGATCACTGACGATCTGGACATCTACACCCGTCTGGG 60

Qy   336 CGGCATGGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCTACCGGCGGTTTCCCGTAG 395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGGCATGGTTGGCGCGCTGACTCCAAAGGCAACTACGCTTCAACCGGCGGTTTCCCGTAG 120

Qy   396 CGAACACGACACTGGGCGTTTCCCCAGTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCG 455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGAACACGACACTGGGCGTTTCCCCAGTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCG 180

Qy   456 TGACATCGCTACACCCGTCTGGAATACCCAGTGGTTAACAACATCGGGACGCGGGCACTGT 515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGACATCGCTACACCCGTCTGGAATACCCAGTGGTTAACAACATCGGGACGCGGGCACTGT 240

Qy   516 GGGTACCGCGTCCTGATAACGGACATGCTGAGCCTGGCGTTCCTACCGCTTCGGTCAGGA 575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GGGTACCGCGTCCTGATAACGGACATGCTGAGCCTGGCGTTCCTACCGCTTCGGTCAGGA 300

Qy   576 AGATGCTGCACCGGTTGTTGCTCCGGCTCCCGGCTCCGGAAGTGGCTACCAAGCA 635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGATGCTGCACCGGTTGTTGCTCCGGCTCCCGGCTCCGGAAGTGGCTACCAAGCA 360

Qy   636 CTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGTACCCTGAAACCGGAAGG 695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAAAGTACCCTGAAACCGGAAGG 420
```

```
Qy   696 TCAGCAGGCTCTGATCAGCTGACAGCTGAGCAACATGATCCGAAAGACGGTTC 755
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCAGCAGGCTCTGATCAGCTGACAGCTGAGCAACATGATCCGAAAGACGGTTC 480

Qy   756 CGCTGTTGTTCTGGCTCACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTC 815
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CGCTGTTGTTCTGGCTCACACCGACCGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTC 540

Qy   816 TGAGAAACGTGCTCAGTCCGTTGTTGACTACCTCGGTTGCTAAAGGCATCCCGGCTGGCAA 875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TGAGAAACGTGCTCAGTCCGTTGTTGACTACCTCGGTTGCTAAAGGCATCCCGGCTGGCAA 600

Qy   876 AATCTCCGCTCGCGCATGGGTGAATCCAACCCGGTTACTGGCAACACTGTGACAACGT 935
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AATCTCCGCTCGCGCATGGGTGAATCCAACCCGGTTACTGGCAACACTGTGACAACGT 660

Qy   936 GAAAGCTCGCGCTGCCCTGATCGATTGCTGGCTCCGGATCGTCGTGTAGAGATCGAAGT 995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GAAAGCTCGCGCTGCCCTGATCGATTGCTGGCTCCGGATCGTCGTGTAGAGATCGAAGT 720

Qy   996 TAAAGG 1001
             ||||||
Db   721 TAAAGG 726
```

```
RESULT 11
US-08-836-500A-3
; Sequence 3, Application US/08836500A
; Patent No. 6197929
; GENERAL INFORMATION:
;     APPLICANT: Binz, Hans
;     APPLICANT: Baussant, Thierry
;     APPLICANT: Haeuw, Jean-Francois
;     APPLICANT: Nguyen Ngoc, Thien
;     TITLE OF INVENTION: Carrier Protein Having an Adjuvant
;     TITLE OF INVENTION: Effect, Immunogenic Complex Containing It, Process for
;     TITLE OF INVENTION: Their Preparation, Nucleotide Sequence and Vaccines
Patent No. 6197929
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
;     STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
;     STREET: 4700
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: U.S.A.
;     ZIP: 60601
; COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/836,500A
;     FILING DATE:
;     CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
;     NAME: Katz, Martin L.
;     REGISTRATION NUMBER: 25,011
;     REFERENCE/DOCKET NUMBER: PIE1514P0180US
; TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312-616-5400
;     TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 537 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;         NAME/KEY: CDS
```

;   LOCATION: 1..537
US-08-836-500A-3

Query Match   51.9%;   Score 537;   DB 3;   Length 537;
Best Local Similarity   100.0%;   Pred. No. 7.5e-143;
Matches 537;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    28  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   87
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   60

Qy    88  ACCGGTTTCTACGGTAACGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT   147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  ACCGGTTTCTACGGTAACGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT   120

Qy   148  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTGACAAATGGGTTAT   207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTGACAAATGGGTTAT   180

Qy   208  GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTCGTTCAAAGCTCAG   267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTCGTTCAAAGCTCAG   240

Qy   268  GGGCGTTCAGTGACCGCTAAACTGGGTTACCCGACTCCAAAGGCAACTACGCTTCTACCGGCGTT   327
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGGCGTTCAGTGACCGCTAAACTGGGTTACCCGACTCCAAAGGCAACTACGCTTCTACCGGCGTT   300

Qy   328  CGTCTCGGGGGGCGGCATGGTTTGGCGCGGCACTGGGCGTTTGCTGGCGCGGCGTAGAGTGGGCT   387
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CGTCTCGGGGGGCGGCATGGTTTGGCGCGGCACTGGGCGTTTGCTGGCGCGGCGTAGAGTGGGCT   360

Qy   388  TCCCGTAGCCGAACAGCGTATTTGCTGGCGCGGCGTTAACAACATCGGCGACGCG   447
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  TCCCGTAGCCGAACAGCGTATTTGCTGGCGCGGCGTTAACAACATCGGCGACGCG   420

Qy   448  GTTACTCGTGACATCGCTACCCGTCTCGGAATACCAGTGGGCGTTAACCCGTTCCTACCGC   507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GTTACTCGTGACATCGCTACCCGTCTCGGAATACCAGTGGGCGTTAACCCGTTCCTACCGC   480

Qy   508  GGCACTGTGGGTACCCGTCCGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  GGCACTGTGGGTACCCGTCCGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   537
```

RESULT 12
US-09-679-750-3
; Sequence 3, Application US/09679750
; Patent No. 6780420
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   Baussant, Thierry
;   Haeuw, Jean-Francois
;   Nguyen Ngoc, Thien
;   TITLE OF INVENTION: Carrier Protein Having an Adjuvant
;   Effect, Immunogenic Complex Containing It, Process for
;   Their Preparation, Nucleotide Sequence and Vaccines
; Patent No. 6780420
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
;   STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
;   4700
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: U.S.A.
;   ZIP: 60601
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/679,750

;   FILING DATE: 08-Oct-2000
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/836,500
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Katz, Martin L.
;   REGISTRATION NUMBER: 25,011
;   REFERENCE/DOCKET NUMBER: PIE1514P0180US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-616-5400
;   TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 537 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..537
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-679-750-3

Query Match   51.9%;   Score 537;   DB 4;   Length 537;
Best Local Similarity   100.0%;   Pred. No. 7.5e-143;
Matches 537;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    28  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   87
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCCCAGTATCACGAC   60

Qy    88  ACCGGTTTCTACGGTAACGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT   147
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  ACCGGTTTCTACGGTAACGTTCCAGAACAACAACGGTCCGACCCGTAACGATCAGCTT   120

Qy   148  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTGACAAATGGGTTAT   207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GGTGCTGGTGCGTTCGGTGGTTACCAGGTTAACCCGTTGACAAATGGGTTAT   180

Qy   208  GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTCGTTCAAAGCTCAG   267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GACTGGCTGGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTCGTTCAAAGCTCAG   240

Qy   268  GGGCGTTCAGTGACCGCTAAACTGGGTTACCCGACTCACTGGACATCTACACC   327
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GGCGTTCAGTGACCGCTAAACTGGGTTACCCGACTCACTGGACATCTACACC   300

Qy   328  CGTCTCGGGGGGCGGCATGGTTTGGCGCGGCACTGGGCGTTTGCTGGCGCGGCGTT   387
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CGTCTCGGGGGGCGGCATGGTTTGGCGCGGCACTGGGCGTTTGCTGGCGCGGCGTT   360

Qy   388  TCCCGTAGCCGAACAGCGTATTTGCTGGCGCGGCGTTAACAACATCGGCGACGCG   447
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  TCCCGTAGCCGAACAGCGTATTTGCTGGCGCGGCGTTAACAACATCGGCGACGCG   420

Qy   448  GTTACTCGTGACATGCTACCCGTCTCGGAATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GTTACTCGTGACATGCTACCCGTCTCGGAATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   480

Qy   508  GGCACTGTGGGTACCCGTCCGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  GGCACTGTGGGTACCCGTCCGATAACGGCATGCTGAGCCTGGGCGTTTCCTACCGC   537
```

RESULT 13
US-09-543-681A-3750
; Sequence 3750, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;   APPLICANT: GARY BRETON
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL:

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3750
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3750

Query Match   44.5%;   Score 460.4; DB 4; Length 1155;
Best Local Similarity   67.1%;   Pred. No. 5.8e-121;
Matches 691; Conservative   0; Mismatches 321; Indels   18; Gaps   2;

Qy   16 GTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGTAAACTGGGTTGGTCC   75
           |||||| | ||| | |||| || | | ||| |||| ||||| |||||||||||| |||
Db   118 GCAGCGCAAGCAGCTCCAAAAGACAATACTTGGTATACCGGTGGTAAATTAGGTTGGTCT 177

Qy   76 CAGTATCACGACACCGGTTTCTACGGTAACGGTTCCAGAACAACAACGGTCCGACCCGT 135
           ||||||| |||
Db   178 CAGTATCAAAGTACTGGTAACAATTGGGACGGCGTGAATATCGGTAACGGTAGCACACAC 237

Qy   136 AACGATCAGCTTGGTGTCGTTGGTGCGTTCGGTGGTGTTACCAGGTTAACCCGTACCTCGGTTTC 195
           | ||| || | || || | | ||| |||| |
Db   238 AAAGACCAAATTGGTGCTGCATTTGCTGGTTATCAGTACACAACCAATACTTAGGTTTT 297

Qy   196 GAAAATGGGTTATGACTGGCTGGCCGTATGGCATATAAAGGCAGCGTTGACAACGGTGCT 255
           |||| | |||||| |||
Db   298 GAACTGGGTTATGACTGGTTAGTCGTATGGCTTATAAAGGTTCATACAACAACGGTGCA 357

Qy   256 TTCAAAGCTCAGGCGTTCAGCTGACCGTCAAACTGGGTTACCCGATCACTGACGATCTG 315
           ||||||| |||
Db   358 TTCAAAGCTCAAGGTATCCAAGTATCCAATTAACCACTAAATTAAGCTATCCAGTAATGGATGACTTA 417

Qy   316 GACATCTACACCCGTCTGGGCGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGCGT 375
           |||||| |||
Db   418 GACGTTTATACACGTTTAGGTGGTGATGGTGCGTGCAGATTCAACTCAACTATTAAT 477

Qy   376 TCTACCGGCGTTTCCCGT----------AGCGAACACGACACTGGCGTTTCCCCAGTA 423
           ||||| |||
Db   478 GCAACTTCAGCAGGAACTCAAAAACGTTTCTCTGAAAATGATACTGGCGTTTCTCCAGTA 537

Qy   424 TTTGCTGCGGCCTAGAGTGGGTGTTTACTCGCTACCCGTCTGGAATACCAG 483
           | |||
Db   538 TTCGCATTAGTTACTGTACTGAATACGCGATCACTCCTAACATTGCTACTCGTCTGTAGTATCAG 597

Qy   484 TGGGTTAACAACATCGGCGACGCGGGACTGTGGTACCCGTCCTGATAACGGCATGGCTG 543
           ||||| |||
Db   598 TGGATCAACAACATCGGTATCGGTGATAAAGGTACTCTGATAAAGCGTCGTCGACAATGTATGCTG 657

Qy   544 AGCCTGGGCGTTCCTACCGGTACCGGTCCGTTCGGTTCAGGAAGATGCTGCACCGGTTGGTTGCTCCGGCT 603
           ||||| |||
Db   658 AGTGTTGGTGTTGCTTACCGGTTTCAACCAAGAAACTCCAGCTCCCAGTTGTTGAGCCAGCT 717

Qy   604 CCG------GCTCCGGCTTCCGGAAGTGCCAGTTGTTGAGAACAAAACCTTTACTCTGCGTTCAGACGTT 657
           ||| |||
Db   718 CCAGTAGTTGCTCCAGCTCCAGTTGTTGAGAACAAAACCTTTACTCTGCGTTCAGACGTT 777

Qy   658 CTGTTCAACTTCAACAAACAAAATCTTCTCTGAAAGCTGAAGGTCAAGAAGCACTGAATGGTCTG 717
           ||||| |||
Db   778 CTGTTCAACTACACAACAAATCTTCTCTGAAAGCTGAAGGTCAAGAAGCACTGAATGGTCTG 837

Qy   718 TACACTCAGCTGACCAACATGGATCCGAAAGACGGTTCCGCTCGTGTTCTGGGCTACACC 777
           ||||| |||
Db   838 TACAATGAATTAGCTAACATCGACCCAACTCAAGGTCGTGTCGTAGTTATCGGTTACACT 897

Qy   778 GACCGCATCGGTTCCGGAAGTTCAAACTACAACAACACCTGCCTTCGTCTGTCAGCAGCAGCGTCAGTCCGTT 837
           ||||| |||
Db   778 GCAGCGTTGAAAACGTTGCATCAACGGTGCAGCAGCTCCAGGGCGTTCAACTCAAACTCGGTT 837

Qy   898 ACCCGATCACTGACGATCTGGACATCTACAACCCGTCTCGGCGGCGATGGTTGGCGCGCTG 957
           ||||| ||||| || |||| ||||||| || | |||| | | | | || ||||
Db   898 GACCGTATCGGTTCTCAAAACTACAACAACACCTGCCTGCTCAAATTACAAGGTGCTGCTAATAAACGTGCTCAATGGTA 957

Qy   838 GTTGACTACCTGGTTGCTAAAGGCATCCCGGCTGCAAAATCTCCGCTCCGGCATGGGT 897
           ||||| ||||||||||||||||| |||| | |||||||||| || || ||||| |||
Db   958 GTTGATTACCTGGTATCTAAAGGTATTCCTGCAAACAGCATCTCTGCAAAGCTCGTGT 1017

Qy   898 GAATCCAACCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATC 957
           |||| ||||| ||||| |||||| ||| |||| || ||||||| | ||||||| |||
Db   1018 AAAGAAACCCAGTAACTGCAACATTTGTGACAACATCAAGCTCGTGTGCAGCTCTGATC 1077

Qy   958 GATTGCCTGGCTCCCGGATCGGTCGTGTAGACGATGAAGTAACAAGAAGTTGTA 1017
           ||||| ||||| |||| ||||| ||| ||||| |||| |||||| |||||||
Db   1078 GATTGCTTAGCGCCAGACGACCGTCGTGTAGAAATCGAAATCCAAGGTACAACTGAAGTTGTT 1137

Qy   1018 ACTCAGCCGG 1027
           ||| |||
Db   1138 GTTCAACCTG 1147

RESULT 14
US-07-794-731A-1
; Sequence 1, Application US/07794731A
; Patent No. 5348867
; GENERAL INFORMATION:
;   APPLICANT: George Georgiou
;   APPLICANT: Joseph A. Francisco
;   APPLICANT: Charles F. Earhart
;   TITLE OF INVENTION: Expression of Proteins on Bacterial
;   TITLE OF INVENTION: Surface
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Arnold, White & Durkee
;   STREET: P.O. Box 4433
;   CITY: Houston
;   STATE: Texas
;   COUNTRY: USA
;   ZIP: 77210
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy Disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WordPerfect 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/794,731A
;   FILING DATE: 19911115
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kitchell, Barbara S.
;   REGISTRATION NUMBER: 33,928
;   REFERENCE/DOCKET NUMBER: GEOG:001
;   TELECOMUNICATION INFORMATION:
;   TELEPHONE: (512) 320-7200
;   TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1273 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-07-794-731A-1

Query Match   21.7%;   Score 224.8; DB 1; Length 1273;
Best Local Similarity   78.0%;   Pred. No. 4.8e-54;
Matches 290; Conservative   0; Mismatches 67; Indels 15; Gaps   1;

Qy   176 TTAACCCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCCGTATGGCATATAAAG 235
           |||||||||||| ||||||| |||||||||| | ||| |||| ||| || |||||
Db   92 TTAACCCGTATGTTGGCTTGAAATGGGTTACGACTGGTCGTATGTCCGTACAAAG 151

Qy   236 GCAGCGGTTGACAACGGTGCTTCCAAAGCTCAGGCGTTCAGCTTCAACTCGGGT 295
           |||||| ||||| ||| || || |||||||| | ||| || ||| |||||||
Db   152 GCAGCGTTGAAAACGGTGCATCAAACGGTGCATCAAACTGGGTT 211

Qy   296 ACCCGATCACTGACGATCTGGACATCTACACCCGTCTCGGCGGCGATGGTTGGCGCGCTG 355
           |||||| ||||| || |||| ||||||| || | |||| | | | | || ||||
Db   296 ACCCGTATCGGTTCTCAAAACTACAACAACACCTGCCTGCTCAAATTACAAGGTGCTGCTG 355

```
Db   212 ACCCAATCACTGACGACCTGACATCTACTCGTCTGGGTGGCATGGTATGGCGTGCAG 271
Qy   356 ACTCCAAAGGCAACTACGCGTTTCTACCGGCGTTCCCGTAGCGAACAGCACCGGCGTTT 415
Db   272 ACACTAAATCCAA--------------CGTTTATGTAAAAACCACGACACCGGCGTTT 316
Qy   416 CCCCAGTATTTGCTGGCGGCGTGAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGG 475
Db   317 CTCCGGTCTTCGCTGGCGGTGTTGAGTACGGCGACCGCGGGCACTCCTGAAATCGCTCGG 376
Qy   476 AATACCAGTGGGTTAACACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACG 535
Db   377 AATACCAGTGGACCAACAACATCGGTGACGGACACACCATCGGCACTCGTCCGGACAACG 436
Qy   536 GCATGCTGAGCC 547
Db   437 GAATTCCGGGTC 448
```

```
Qy   296 ACCCGATCACTGACGACTCGGACACTTCACCCGTCTGGGCGGGCATGGTTTGGCGCCTG 355
Db   212 ACCCAATCACTGACGACCTGACATCTACTCGTCTGGGTGGTATGGCGTGCAG 271
Qy   356 ACTCCAAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACAGCACCGGCGTTT 415
Db   272 --------CGTTTATGTAAAAACCACGACACCGGCGTTT 316
Qy   416 CCCCAGTATTTGCTGGCGGCGTGAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGG 475
Db   317 CTCCGGTCTTCGCTGGCGGTGTTGAGTACGGCGACCGCGGGCACTCCTGAAATCGCTCGG 376
Qy   476 AATACCAGTGGGTTAACACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACG 535
Db   377 AATACCAGTGGACCAACAACATCGGTGACGGACACACCATCGGCACTCGTCCGGACAACG 436
Qy   536 GCATGCTGAGCC 547
Db   437 GAATTCCGGGTC 448
```

Search completed: January 19, 2005, 16:38:45
Job time : 122 secs

RESULT 15
US-07-794-731A-2
; Sequence 2, Application US//07794731A
; Patent No. 5348867
; GENERAL INFORMATION:
; APPLICANT: George Georgiou
; APPLICANT: Joseph A. Francisco
; APPLICANT: Charles F. Earhart
; TITLE OF INVENTION: Expression of Proteins on Bacterial
; TITLE OF INVENTION: Surface
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/794,731A
; FILING DATE: 19911115
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: GEOG:001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 320-7200
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1273 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-794-731A-2

Query Match   21.7%;   Score 224.8; DB 1; Length 1273;
Best Local Similarity   78.0%;   Pred. No. 4.8e-54;
Matches 290; Conservative   0; Mismatches 67; Indels 15; Gaps   1;

```
Qy   176 TTAACCCGTACCTCCGGTTTCGAAATGGGTTATGACTGGCTGGGCCCGTATGCGCATATAAAG 235
Db    92 TTAACCCGTATGTTGGCTTTGAAATGGGTTACGACTGGTTAGTGCGTATGCCGTACAAAG 151
Qy   236 GCAGCCGTTGACAACGGTGCTTCAAAGTCAGGGCGTTCAGGCTGCTAAACTGGGTT 295
Db   152 GCAGCCGTTGAAAACGGTGCATACAAAGCTCAGGGCCGTTCAACTGACCGCTAAACTGGGTT 211
```

^

,

```
                OM nucleic - nucleic search, using sw model

                Copyright (c) 1993 - 2005 . Compugen Ltd.

                        GenCore version 5.1.6

Run on:          January 19, 2005, 10:15:18 ; Search time 603 Seconds
                                            (without alignments)
                                9010.200 Million cell updates/sec

Title:           US-09-913-772-1
Perfect score:   1035
Sequence:        1 atgaaagcaattccgtact.........taactcagccggcgggtcaa 1035

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters:   8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N Geneseq_23Sep04:*
                  1: geneseqn1980s:*
                  2: geneseqn1990s:*
                  3: geneseqn2000s:*
                  4: geneseqn2001as:*
                  5: geneseqn2001bs:*
                  6: geneseqn2002as:*
                  7: geneseqn2002bs:*
                  8: geneseqn2003as:*
                  9: geneseqn2003bs:*
                 10: geneseqn2003cs:*
                 11: geneseqn2003ds:*
                 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1035	100.0	1035	3	AAA96568	Aaa96568 DNA encod
2	1035	100.0	1035	3	AAA63961	Aaa63961 DNA encod
3	1035	100.0	1035	3	AAA15498	Aaa15498 DNA encod
4	1035	100.0	1035	3	AAA75036	Aaa75036 cDNA enco
5	1035	100.0	1035	3	AAA75881	Aaa75881 DNA encod
6	1035	100.0	1035	3	AAA63956	Aaa63956 DNA encod
7	1035	100.0	1035	4	AAH74731	Aah74731 Nucleotid
8	1032	99.7	1032	4	AAF80152	Aaf80152 Nucleotid
9	1030.2	99.5	1035	2	AAV13868	Aav13868 Gene codi
10	1026.2	99.1	1356	4	AAF80153	Aaf80153 Nucleotid
11	1024	98.9	1032	2	AAZ30477	Aaz30477 K.pneumon
12	1024	98.9	1032	4	AAH78461	Aah78461 Nucleotid
13	1024	98.9	1032	5	AAF90077	Aaf90077 Nucleotid
14	1009.8	97.6	1158	11	ACH94485	Ach94485 Klebsiell
15	1009.8	97.6	1251	11	ACH94572	Ach94572 Klebsiell
16	1008	97.4	1008	2	AAV13867	Aav13867 Gene codi
17	1007	97.3	1008	2	AAT31607	Aat31607 Klebsiell
18	1006.4	97.2	1008	2	AAT03490	Aat03490 K.pneumon
19	724.4	70.0	730	4	AAT28425	Aat28425 K. pneumo
20	724.4	70.0	730	4	ABA76835	Aba76835 Klebsiell
21	681	65.8	5579	2	AAQ02032	Aaq02032 Plasmid p

22	642.4	62.1	1018	5	ADM19682	Adm19682 Novel hum
23	565	54.6	567	2	AAV13869	Aav13869 Gene codi
24	537	51.9	537	2	AAT31608	Aat31608 Klebsiell
25	486	47.0	6477	2	AAQ02030	Aaq02030 Plasmid p
26	460.4	44.5	1155	10	ADF03465	Adf03465 Bacterial
27	457.6	44.2	1107	10	ACF68624	Acf68624 Photorhab
28	457.6	44.2	110000	10	ACF67367_13	Continuation (14 o
29	457.6	44.2	182624	10	ACF65379_13	Acf65379 Photorhab
30	403	38.9	411	6	ABN83312	Abn83312 Partial O
31	303.4	29.3	3630	6	AAD40180	Aad40180 pDomp pla
32	300.6	29.0	3600	6	AAD40168	Aad40168 pEomp pla
33	298.6	28.9	3660	6	AAD40168	Aad40168 pComp pla
34	298.6	28.9	4543	6	AAD40198	Aad40198 pOHL44a p
35	224.8	21.7	1273	2	AAQ42410	Aaq42410 Tripartit
36	219	21.2	3817	10	ADD14893	Add14893 Phage dis
37	216	20.9	216	2	AAT31609	Aat31609 Klebsiell
38	161	15.6	1095	3	AAC79663	Aac79663 Virulence
39	161	15.6	1095	5	ABQ83540	Abq83540 Actinobac
40	159.4	15.4	1922	3	AAA38557	Aaa38557 Actinobac
41	159	15.4	159	2	AAT31610	Aat31610 Klebsiell
42	154.8	15.0	690	10	ACF65985	Acf65985 Photorhab
43	135	13.0	1035	6	ABA91422	Aba91422 Haemophil
44	134.2	13.0	1728	5	AAS82075	Aas82075 DNA encod
45	134.2	13.0	1728	5	AAS82046	Aas82046 DNA encod

```
                        ALIGNMENTS

RESULT 1
AAA96568
ID   AAA96568 standard; DNA; 1035 BP.
XX
AC   AAA96568;
XX
DT   08-FEB-2001  (first entry)
XX
DE   DNA encoding a P40 polypeptide of Klebsiella pneumoniae.
XX
KW   P40; membrane fraction; Gram-negative bacteria; anticancer;
KW   immune response; mononuclear blood cell; tumour necrosis factor-alpha;
KW   interleukin-12; antitumour; cancer; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1035
FT                  /*tag= a
FT                  /product= "P40"
XX
PN   WO200054790-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-FR000623.
XX
PR   15-MAR-1999;  99FR-00003154.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Libon C,  Corvaia N,  Beck A,  Bonnefoy J;
XX
DR   WPI; 2000-587477/55.
DR   P-PSDB; AAB18994.
XX
PT   Use of membrane fractions from Gram-negative bacteria as immunostimulants
PT   for the treatment or prevention of cancer, increases effect of e.g.
PT   chemotherapeutic agents.
XX
PS   Disclosure; Page 25-27; 34pp; French.
XX
CC   The present sequence encodes a P40 polypeptide of Klebsiella pneumoniae.
CC   The protein is found in the membrane fraction, and is used in the method
```

```
CC    of the invention. The specification describes the use of a membrane
CC    fraction from Gram-negative bacteria for the preparation of an
CC    immunostimulating composition that can induce an anticancer immune
CC    response. The membrane fraction not only stimulates proliferation of
CC    human mononuclear blood cells (immunostimulation) but also induces
CC    production of tumour necrosis factor-alpha and interleukin (IL)-12, which
CC    are known to have antitumour activity, so that it improves the effects of
CC    other co-administered anticancer treatments (chemotherapy or radiation).
CC    The membrane fraction is used for treatment and prevention of cancer
CC    (particularly of the bladder, prostate, colon or liver) and also
CC    malignant melanomas
XX
SQ    Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1035;  DB 3;  Length 1035;
Best Local Similarity   100.0%;  Pred. No. 3.5e-283;
Matches 1035;        Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy         61 AAACTGGGTTGGTCCCCAGTATCACGACACCGGTTTCTACCGGTAACGGTTTCCAGAACAAC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAACTGGGTTGGTCCCCAGTATCACGACACCGGTTTCTACCGGTAACGGTTTCCAGAACAAC 120

Qy        121 AACGGTCCGACCCGTAACGATCAGCTTGGTGTCGGTGCTGTTAACCAGTTAAC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACGGTCCGACCCGTAACGATCAGCTTGGTGTCGGTGCTGTTAACCAGTTAAC 180

Qy        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy        241 GTTGACAACGGTGCTTCAAAGCTCAGCTGGCCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTGACAACGGTGCTTCAAAGCTCAGCTGGCCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy        301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCATGGTTGGGCGCTGACTCC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCATGGTTGGGCGCTGACTCC 360

Qy        361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACGACGACTGGCGTTCCCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACGACGACTGGCGTTCCCCA 420

Qy        421 GTATTGCTGCGGCGGCCTAGAGTGGGCGTGTTACTCGTGACATCGTCACCCGTCTGGAATAC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTATTGCTGCGGCGGCCTAGAGTGGGCGTGTTACTCGTGACATCGTCACCCGTCTGGAATAC 480

Qy        481 CAGTGGTTAACAACATCGGCGACCGGCGGCACTGTGGTACCCGTCCTGATAACGGCATG 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGTGGTTAACAACATCGGCGACCGGCGGCACTGTGGTACCCGTCCTGATAACGGCATG 540

Qy        541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600

Qy        601 GCTCCCGGCTCCGGCGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTCCCGGCTCCGGCGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy        661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy        721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780

Qy        781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy        841 GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAAATCTCCGCTCCGGCATGGTGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAAATCTCCGCTCCGGCATGGTGAA 900

Qy        901 TCCAACCGGTTACTGGCACACCTGTGACAACGTGAAAGCTCGGCGCTGCCCTGATCGAT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCCAACCGGTTACTGGCACACCTGTGACAACGTGAAAGCTCGGCGCTGCCCTGATCGAT 960

Qy        961 TGCCTGGCTCCGGATCGTGCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCCTGGCTCCGGATCGTGCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy       1021 CAGCCGCGGGTTAA 1035
              ||||||||||||||
Db       1021 CAGCCGCGGGTTAA 1035


RESULT 2
AAA63961
ID   AAA63961 standard; DNA; 1035 BP.
XX
AC   AAA63961;
XX
DT   04-DEC-2000  (first entry)
XX
DE   DNA encoding an outer membrane protein A (OmpA), designated P40.
XX
KW   Outer membrane protein A; OmpA; P40; cytotoxic T cell response;
KW   CTL response; tumour cell; vaccine; infection; tumour; melanoma;
KW   genetic vaccine; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1035
FT                   /*tag= a
FT                   /product= "outer membrane protein A (OmpA) P40"
XX
PN   WO20004628-A1.
XX
PD   24-AUG-2000.
XX
PF   17-FEB-2000; 2000WO-FR000393.
XX
PR   17-FEB-1999;  99FR-00001917.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Renno T, Bonnefoy J;
XX
DR   WPI; 2000-543667/49.
DR   P-PSDB; AAB08317.
XX
PT   Use of enterobacterial outer membrane protein A in vaccines for inducing
PT   cytotoxic T cell responses, useful for treating or preventing infections
PT   and tumors.
XX
PS   Claim 24; Page 36-37; 45pp; French.
XX
CC   The present sequence encodes a Klebsiella pneumoniae outer membrane
CC   protein A (OmpA), designated P40. The enterobacterial OmpA polypeptide,
CC   or its fragments, is used for preparing a composition that increases, or
CC   increases, the cytotoxic T cell (CTL) response against an infectious
CC   agent or tumour cell. Compositions containing OmpA, optionally mixed with
CC   or coupled to a suitable antigen or hapten, are used as vaccines for
CC   treatment or prevention of infections caused by viruses, bacteria, fungi
CC   and parasites or tumors, particularly where associated with an antigen
CC   and specifically melanoma. Nucleic acids that encode OmpA (or its fusion
CC   with antigens or haptens) are useful as genetic vaccines again for
CC   treating infections and tumors
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;
```

```
Qy        1021 CAGCCGGCGGGGTTAA 1035
               ||||||||||||||||
Db        1021 CAGCCGGCGGGGTTAA 1035
```

RESULT 3
AAA15498
ID   AAA15498 standard; DNA; 1035 BP.
XX
AC   AAA15498;
XX
DT   04-SEP-2000 (first entry)
XX
DE   DNA encoding a Klebsiella P40 protein.
XX
KW   P40 protein; outer membrane protein A; OmpA; antigen-presenting cell;
KW   dendritic cell; antigen delivery; immune response; cancer;
KW   tumour-associated antigen; autoimmune disease; allergy; graft rejection;
KW   cardiovascular disease; central nervous system disease; inflammation;
KW   infection; immune deficiency; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1035
FT   /*tag= a
FT   /product= "P40 protein"
XX
PN   WO20027432-A1.
XX
PD   18-MAY-2000.
XX
PF   08-NOV-1999; 99WO-FR002734.
XX
PR   06-NOV-1998; 98FR-0014007.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI   Bonnefoy J, Lecoanet S, Aubry J, Jeannin P, Baussant T;
XX
DR   WPI; 2000-387342/33.
DR   P-PSDB; AAY93341.
XX
PT   Use of enterobacterial outer membrane protein A for delivering active
PT   substances, particularly immunogens for treating or preventing e.g.
PT   cancer, to antigen presenting cells.
XX
PS   Example 1; Page 26-28; 35pp; French.
XX
CC   The present sequence encodes a P40 protein. The protein is an outer
CC   membrane protein A (OmpA). The protein is used in pharmaceutical
CC   compositions for specific targeting of an active substance to antigen-
CC   presenting cells (APCs), especially dendritic cells. OmpA binds
CC   specifically to APCs and is internalised by them (in contrast to other
CC   protein carriers such as tetanus toxoid). The OmpA protein is used to
CC   deliver an antigen or hapten to modify (specifically to improve) an
CC   immune response, especially for treatment or prevention of cancers
CC   (particularly those that express a associated-associated antigen),
CC   autoimmune disease, allergy, graft rejection, cardiovascular or central
CC   nervous system diseases, inflammation, infection or immune deficiency
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;
```

```
Query Match     100.0%;  Score 1035;  DB 3;  Length 1035;
Best Local Similarity  100.0%;  Pred. No. 3.5e-283;
Matches 1035;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy       1 ATGAAAGACAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGAAAGACAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy      61 AAACTGGGTGGTCCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGACAAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAACTGGGTGGTCCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGACAAC 120
```

```
Query Match     100.0%;  Score 1035;  DB 3;  Length 1035;
Best Local Similarity  100.0%;  Pred. No. 3.5e-283;
Matches 1035;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy       1 ATGAAAGACAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGAAAGACAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy      61 AAACTGGGTGGTCCGACCCGTAACGATCAGCTGGTGGTGTCGTTCGGTGGTGTGGTTAAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAACTGGGTGGTCCGACCCGTAACGATCAGCTGGTGGTGTCGTTCGGTGGTGTGGTTAAC 120

Qy     121 AACGGTCCGACCCGTAACGATCAGCTGGTGGTGTCGTTCGGTGGTGTGGTTACCAGGTTAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AACGGTCCGACCCGTAACGATCAGCTGGTGGTGTCGTTCGGTGGTGTGGTTACCAGGTTAAC 180

Qy     181 CCGTACCCGGTTTCGAAATGGGTTATGACTGCTGGCCGTATGGCATATAAAGGCAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCGTACCCGGTTTCGAAATGGGTTATGACTGCTGGCCGTATGGCATATAAAGGCAGC 240

Qy     241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy     301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCGCATGTTTGGCGGCCTGACTCC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCGCATGTTTGGCGGCCTGACTCC 360

Qy     361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420

Qy     421 GTATTTGCTGGCCGGCGTAGAGTGGGCTGTTACTCGTGACATCCGTACCGTCTGGAATAC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTATTTGCTGGCCGGCGTAGAGTGGGCTGTTACTCGTGACATCCGTACCGTCTGGAATAC 480

Qy     481 CAGTGGGTAACACATCGGGCGACGCGGGCACTGTGGTACCCGTCCTGATAACGGCATG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CAGTGGGTAACACATCGGGCGACGCGGGCACTGTGGTACCCGTCCTGATAACGGCATG 540

Qy     541 CTGAGCCTGGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCGCACCGGTTGTTGCTCCG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CTGAGCCTGGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCGCACCGGTTGTTGCTCCG 600

Qy     601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy     661 TTCAACTTCAACAAAGTACCCTGAAACCGGAAGGTCAGCGCTGGATCAGCTGTAC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TTCAACTTCAACAAAGTACCCTGAAACCGGAAGGTCAGCGCTGGATCAGCTGTAC 720

Qy     721 ACTCAGCTGAGCAACAATGATGATCCGAAAGACGGTTCCCGTGTTGTTCGGCTACACCGAC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACTCAGCTGAGCAACAATGATGATCCGAAAGACGGTTCCCGTGTTGTTCGGCTACACCGAC 780

Qy     781 CCGATCCGGTTCCGAAGCTTACAACAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CCGATCCGGTTCCGAAGCTTACAACAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy     841 GACTACCTGGTTGCTAAAGGCATCCCGCGCTGCAAAATCTCCGCGTCGCGGCATGGGTGAA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACTACCTGGTTGCTAAAGGCATCCCGCGCTGCAAAATCTCCGCGTCGCGGCATGGGTGAA 900

Qy     901 TCCAACCCGGTTACTGCGCAACACCTGTGACAACGTGAAAGCTCGCGCCCTGATCGAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TCCAACCCGGTTACTGCGCAACACCTGTGACAACGTGAAAGCTCGCGCCCTGATCGAT 960

Qy     961 TGCCTGGCTCCGGATCGTGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TGCCTGGCTCCGGATCGTGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
```

DT   02-JAN-2001 (first entry)
XX
DE   cDNA encoding a P40 polypeptide (an outer membrane protein A (OmpA)).
XX
KW   P40; outer membrane protein A; OmpA; immunogen; cytokine; growth factor;
KW   hormone; tumour-specific marker; vaccine; cancer; contraceptive; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1035
FT   /*tag= a
FT   /product= "P40"
XX
PN   FR2789902-A1.
XX
PD   25-AUG-2000.
XX
PF   24-FEB-1999; 99FR-00002314.
XX
PR   24-FEB-1999; 99FR-00002314.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI   Goetsch L, Corvaia N, Beck A, Haeuw JF, Bonnefoy JY;
XX
DR   WPI; 2000-573921/54.
DR   P-PSDB; AAB08825.
XX
PT   Use of enterobacterial outer membrane protein as immunogenic carrier,
PT   particularly for contraceptive and anti-cancer vaccines, provides strong
PT   humoral response.
XX
PS   Claim 14; Page 21-22; 34pp; French.
XX
CC   The present sequence encodes a P40 polypeptide of Klebsiella pneumoniae.
CC   P40 is an enterobacterial outer membrane protein A (OmpA). It can be
CC   associated an immunogen, and used to prepare a pharmaceutical composition
CC   for improving the immunological response to the immunogen. The immunogen
CC   is selected from cytokines, growth factors or hormones (or their
CC   receptors) and/or tumour-specific markers. Compositions containing OmpA
CC   induce a strong and specific antibody response. The compositions of the
CC   invention are especially useful in vaccines to prevent or treat cancer or
CC   as contraceptives
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;

Query Match   100.0%; Score 1035; DB 3; Length 1035;
Best Local Similarity   100.0%; Pred. No. 3.5e-283;
Matches 1035; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGAACAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGGTTTCCAGGTTAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGGTTTCCAGGTTAAC 180

Qy   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCATATAAAGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCATATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTCAAAGCTCAGCGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTCAAAGCTCAGCGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy   301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCCATGGTTTGGGCGCTGACTCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCCATGGTTTGGGCGCTGACTCC 360

RESULT 4
AAA75036
ID   AAA75036 standard; cDNA; 1035 BP.
XX
AC   AAA75036;
XX

XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-FR000622.
XX
PR   15-MAR-1999; 99FR-00003153.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE..
XX
PI   Libon C, Corvaia N, N'guyen TN, Beck A, Bonnefoy J;
XX
DR   WPI; 2000-587476/55.
DR   P-PSDB; AAB18804.
XX
PT   Use of Klebsiella membrane fraction as adjuvant, for e.g. antitumor or
PT   antiviral vaccines, to direct a Th1, or mixed, immune response against
PT   associated antigen.
XX
PS   Disclosure; Page 27-29; 36pp; French.
XX
CC   The present sequence encodes a Klebsiella pneumoniae P40 polypeptide. The
CC   protein is isolated from a membrane fraction. The specification describes
CC   the use of a membrane fraction from Klebsiella pneumoniae, associated
CC   with an antigen or hapten, for preparation of a pharmaceutical
CC   composition that directs a Th1, or mixed Th1/Th2 immune response. The
CC   composition is used for treatment or prevention of infectious diseases
CC   (viral, bacterial, fungal or parasitic) or cancers, most especially
CC   infections by paramyxoviruses, specifically respiratory syncytial virus
CC   or parainfluenza
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1035;   DB 3;   Length 1035;
Best Local Similarity   100.0%;   Pred. No. 3.5e-283;
Matches 1035;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy  121 AACGGTCCGACCCGTAACGATCAGCAGCTTGGTGTGCGTTCGGTGCGTTCGGTGGTTAAC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACGGTCCGACCCGTAACGATCAGCAGCTTGGTGTGCGTTCGGTGCGTTCGGTGGTTAAC 180

Qy  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy  241 GTTGACAACGGTGCTCTTTCAAAGCTCAGGCGTTCAGCTGACCGGTTCAGCTGACCCCG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTTGACAACGGTGCTCTTTCAAAGCTCAGGCGTTCAGCTGACCGGTTCAGCTGACCCCG 300

Qy  301 ATCACTGACGACGATCTGGACATCTACACCGGCGTTCTCGGGCGGCCGCTGACTCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATCACTGACGACGATCTGGACATCTACACCGGCGTTCTCGGGCGGCCGCTGACTCC 360

Qy  361 AAAGGCAACTACCGCTTCTACCGGCGTTTCCCGTAGCGGAACACGACACTGGCGTTCCCCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGGCAACTACCGCTTCTACCGGCGTTTCCCGTAGCGGAACACGACACTGGCGTTCCCCA 420

Qy  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTCGTAACCCGTCTGGAATAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTCGTAACCCGTCTGGAATAC 480

Qy  481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGCTACCCGTCTGATAACGGCATG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGCTACCCGTCTGATAACGGCATG 540

Qy  541 CTGAGCCTGGGCGGCGTTTCCTCACGCTTCGGTTCGGTCCGGAAGATGCTGCAGGAGATGGCTGCTCCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTGAGCCTGGGCGGCGTTTCCTCACGCTTCGGTTCGGTCCGGAAGATGCTGCAGGAGATGGCTGCTCCG 600

Qy  601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGCTCGACGTTCTG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGCTCGACGTTCTG 660

Qy  661 TTCAACTTCAACAAAGTACCCTGAAACCGGAAGGTCAGCGGCTCTTGGATCAGCAGCTGTAC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TTCAACTTCAACAAAGTACCCTGAAACCGGAAGGTCAGCGGCTCTTGGATCAGCAGCTGTAC 720

Qy  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTTGTTCTTGGCTACACCGAC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTTGTTCTTGGCTACACCGAC 780

Qy  781 CGCATCGGTTCCGAAGCTTACAACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CGCATCGGTTCCGAAGCTTACAACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy  841 GACTACCTGGTTGCTAAAGGCATCCCGCTGGCAAAATCTCCGCTCCGCGGCATGGGTGAA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACTACCTGGTTGCTAAAGGCATCCCGCTGGCAAAATCTCCGCTCCGCGGCATGGGTGAA 900

Qy  901 TCCAACCCGGTTACTGGCAACACGTTGACACGTGAAAGCTGAAAGCTGCCCTGATCGAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCAACCCGGTTACTGGCAACACGTTGACACGTGAAAGCTGAAAGCTGCCCTGATCGAT 960

Qy  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy 1021 CAGCCGCGGGTTAA 1035
          ||||||||||||||
Db 1021 CAGCCGCGGGTTAA 1035
```

RESULT 5
AAA75881
ID   AAA75881 standard; DNA; 1035 BP.
XX
AC   AAA75881;
XX
DT   22-JAN-2001 (first entry)
XX
DE   DNA encoding a Klebsiella pneumoniae P40 polypeptide.
XX
KW   P40 polypeptide; membrane fraction; antigen; hapten; immune response;
KW   infectious disease; cancer; paramyxovirus infection;
KW   respiratory syncytial virus; parainfluenza; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1035
FT   /*tag= a
FT   /product= "P40 polypeptide"
XX
PN   WO200054789-A1.

PT   cytotoxic lymphocyte response.
XX
PS   Example 1; Page 28-29; 35pp; French.
XX
CC   The present sequence encodes a Klebsiella pneumoniae outer membrane
CC   protein A (OmpA), designated P40. The enterobacterial OmpA polypeptide,
CC   or its fragments, is used for preparing a composition that induces, or
CC   increases, the cytotoxic T cell (CTL) response against tumour cells.
CC   Compositions containing OmpA, optionally mixed with or coupled to a
CC   suitable antigen or hapten, are used as vaccines for treatment or
CC   prevention of tumors, particularly where associated with an antigen and
CC   specifically melanoma. Nucleic acids that encode OmpA (or its fusion with
CC   antigens or haptens) are useful as genetic vaccines again for treating
CC   tumors
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;

Query Match   100.0%;   Score 1035; DB 3; Length 1035;
Best Local Similarity 100.0%;   Pred. No. 3.5e-283;
Matches 1035;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy    61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAAGGTTTCCAGAACAAC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAAGGTTTCCAGAACAAC 120

Qy   121 AACGGTTCCGACCCGTAACGATCAGCTTGGTGTCGTGGTCGTTCGGTGGTTACCAGGTTAAC 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTTCCGACCCGTAACGATCAGCTTGGTGTCGTGGTCGTTCGGTGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTCGGTTTCGAAATGGTTATGACTGGTGGCCGTGGGCCGTGTATGGCATATAAAGGCAGC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTTCGAAATGGTTATGACTGGTGGCCGTGGGCCGTGTATGGCATATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTCAAAGCTTCAAAGCTTCAGCTGACGGGTTCAGGGGTTCAGCTGACCGTTACCCG 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTCAAAGCTTCAAAGCTTCAGCTGACGGGTTCAGGGGTTCAGCTGACCGTTACCCG 300

Qy   301 ATCACTGACGATCTGACATCTACACGCCGGTTTCCGGCGGCGCATGGTTTGGCGCGCTGACTCC 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGACATCTACACGCCGGTTTCCGGCGGCGCATGGTTTGGCGCGCTGACTCC 360

Qy   361 AAAGGCAACTACGCTTCTACCGGCGGTTTCCCGTAGGCGAACACGACACTGGCGTTCCCCA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGGTTTCCCGTAGGCGAACACGACACTGGCGTTCCCCA 420

Qy   421 GTATTTGCTGCGGCGGGCTAGAGTGGCGCTGTTACTCGTGACATCGCTACCCCGTCTGAATAC 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGCGGCGGGCTAGAGTGGCGCTGTTACTCGTGACATCGCTACCCCGTCTGGAATAC 480

Qy   481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGCACTGTGGGTACCCGCATG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGCACTGTGGGTACCCGCATG 540

Qy   541 CTGAGCCTGGGCGGTTCCTACCGCGTTGTTGCTCCG 600
         ||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGGTTCCTACCGCGTTGTTGCTCCG 600

Qy   601 GCTCCGGCTCCGGCGGAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGGCTCCGGCGGAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy   661 TTCAACTTCAACAAAGCTACCCCTGAAACCGGAGGCTCTGGATCAGCTGTAC 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGCTACCCCTGAAACCGGAGGCTCTGGATCAGCTGTAC 720

Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780

Qy   781 CGCATCGGTGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATCGGTGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCCGCGCGCATGGGTGAA 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCCGCGCGCATGGGTGAA 900

Qy   901 TCCAACCCGGTTACTGGCAACACTGTGACAACGTGAAAGTCGAAAGCTGCCCTGATCGAT 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGGCAACACTGTGACAACGTGAAAGTCGAAAGCTGCCCTGATCGAT 960

Qy   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGAGTTGTAACT 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGAGTTGTAACT 1020

Qy  1021 CAGCCGGCGGGGTTAA 1035
         ||||||||||||||||
Db  1021 CAGCCGGCGGGGTTAA 1035
```

RESULT 6
AAA63956
ID   AAA63956 standard; DNA; 1035 BP.
XX
AC   AAA63956;
XX
DT   04-DEC-2000 (first entry)
XX
DE   DNA encoding an outer membrane protein A (OmpA), designated P40.
XX
KW   Outer membrane protein A; OmpA; P40; cytotoxic T cell response; tumour;
KW   CTL response; tumour cell; vaccine; melanoma; genetic vaccine; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1035
FT   /tag= a
FT   /product= "outer membrane protein A (OmpA) P40"
XX
PN   WO20004862 9-A1.
XX
PD   24-AUG-2000.
XX
PF   17-FEB-2000; 2000WO-FR000394.
XX
PR   17-FEB-1999; 99FR-00001917.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI   Renno T, Romero P, Miconnet I, Carottini J, Bonnefoy J;
XX
DR   WPI; 2000-549238/50.
DR   P-PSDB; AAB08341.
XX
PT   Use of enterobacterial outer membrane protein A in vaccines, used to
PT   treat or prevent melanoma, includes melanoma-specific peptide and induces

```
CC   human immune deficiency virus, respiratory syncytial virus, measles,
CC   mumps, tuberculosis etc.), but also against fungi, parasites, autoimmune
CC   diseases, graft rejection, cardiovascular disease, inflammation and
CC   immune deficiency
XX
SQ   Sequence 1035 BP; 227 A; 283 C; 288 G; 237 T; 0 U; 0 Other;

Query Match         100.0%;  Score 1035;  DB 4;  Length 1035;
Best Local Similarity   100.0%;  Pred. No. 3.5e-283;
Matches 1035;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy     61 AAACTGGGTTGGTCCGTATCACGACACCGTTTCACGTAACGGTTTCAGAACAAC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAACTGGGTTGGTCCGTATCACGACACCGTTTCACGTAACGGTTTCAGAACAAC 120

Qy    121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCGTTCGGTGGTGGTTACCAGGTTAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCGTTCGGTGGTGGTTACCAGGTTAAC 180

Qy    181 CCGTACCTCGGTTCGAAATGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGTACCTCGGTTCGAAATGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy    241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGTTCAGCTGACCGGCGCTAAACTGGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGTTCAGCTGACCGGCGCTAAACTGGGTTACCCG 300

Qy    301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGTTGACTCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGTTGACTCC 360

Qy    361 AAAGGCAACTACGCTTCACCGGCCGTTCCCGTAGCGAACACGACACACTGGCGTTTCCCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGGCAACTACGCTTCACCGGCCGTTCCCGTAGCGAACACGACACACTGGCGTTTCCCCA 420

Qy    421 GTATTGCTGGCGGCGTAGAGTGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATTGCTGGCGGCGTAGAGTGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCGTCCGTCCTGATAACGGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCGTCCGTCCTGATAACGGCATG 540

Qy    541 CTGAGCCTGGGGCGTTTCCTACCCGTTCGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGAGCCTGGGGCGTTTCCTACCCGTTCGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600

Qy    601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCCGAAGTCTGACGTTCTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCCGAAGTCTGACGTTCTG 660

Qy    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGTCAGCAGCTGTAC 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGTCAGCAGCTGTAC 720

Qy    721 ACTCAGCTGAGCAACATGATCGATCCGCTACACCGAC 780
          |||||||||||||||||||||||||||||||||||||
Db    721 ACTCAGCTGAGCAACATGATCGATCCGCTACACCGAC 780

Qy    781 CGATCCGGTTCCGAAGCTTCACACCAGCAGTCGTCTGAGAAACGTGCTCGTTGTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGATCCGGTTCCGAAGCTTCACACCAGCAGTCGTCTGAGAAACGTGCTCGTTGTT 840

Qy    781 CGCATCCGGTTCCGAAGCTTACACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGCATCCGGTTCCGAAGCTTACACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy    841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGCAAAATCTCCGCTCGCGGGTGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGCAAAATCTCCGCTCGCGGGTGAA 900

Qy    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTGAAGTCCGTGCCCTGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTGAAGTCCGTGCCCTGAT 960

Qy    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGCGGGGTTAA 1035
          |||||||||||||||
Db   1021 CAGCCGCGGGGTTAA 1035
```

```
RESULT 7
AAH74731
ID   AAH74731 standard; cDNA; 1035 BP.
XX
AC   AAH74731;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of an outer membrane protein A, P40.
XX
KW   Outer membrane protein A; P40; antigen presenting cell; vaccine;
KW   antiviral; antibacterial; anticancer; autoimmune disease; inflammation;
KW   graft rejection; cardiovascular disease; immune deficiency; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..1035
FT                    /*tag= a
FT                    /product= "P40"
XX
PN   FR2803302-A1.
XX
PD   06-JUL-2001.
XX
PF   04-JAN-2000; 2000FR-00000070.
XX
PR   04-JAN-2000; 2000FR-00000070.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Baussant T, Jeannin P, Delneste Y, Lawny F, Bonnefoy JY;
XX
DR   WPI; 2001-427232/46.
DR   P-PSDB; AAG63698.
XX
PT   Preparing purified polypeptide soluble in absence of detergent, useful
PT   for modulating the immune system, e.g. in vaccines, by removal of
PT   detergent, denaturing and molecular sieving.
XX
PS   Example 1; Page 22-24; 34pp; French.
XX
CC   The present sequence encodes an outer membrane protein A (P40) of
CC   Klebsiella pneumoniae. The protein is soluble in aqueous solvent in
CC   absence of detergent. The specification describes a method for the
CC   preparation of this polypeptide. The P40 protein binds selectively to
CC   antigen-presenting cell, so provides targeting, proliferation and/or
CC   expression of molecules by these cells. P40 is used, alone or as an
CC   adjuvant, to produce therapeutic compositions that are soluble in absence
CC   of detergent, especially when formulated with an antigen or hapten for
CC   modulating the host's immune system. Especially, it is used to prepare
CC   vaccines, especially antiviral, antibacterial or anticancer (e.g. against
```

```
Db    901  TCCAACCCGGTTACTGCCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATCGAT  960

Qy    961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020

Qy   1021  CAGCCGCGGGGTTAA  1035
           |||||||||||||||
Db   1021  CAGCCGCGGGGTTAA  1035
```

RESULT 8
AAF80152
ID   AAF80152 standard; DNA; 1032 BP.
XX
AC   AAF80152;
XX
DT   11-JUN-2001 (first entry)
XX
DE   Nucleotide sequence of an outer membrane protein A (OmpA) P40.
XX
KW   Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;
KW   respiratory syncytial virus; RSV; RSV infection; lung; respiratory tract;
KW   vaccine; ss.
XX
OS   Klebsiella pneumoniae.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1032
FT   /*tag= a
FT   /product= "outer membrane protein A (OmpA) P40"
FT   /note= "no termination codon"
XX
PN   WO200121203-A1.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-FR002626.
XX
PR   23-SEP-1999; 99FR-00011888.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI   Corvaiea N, Goestch L;
XX
DR   WPI; 2001-257929/26.
DR   P-PSDB; AAB67770.
XX
PT   Vaccine against respiratory syncytial virus, comprises enterobacterial
PT   outer membrane protein and viral immunogen, provides protective response
PT   throughout the respiratory tract.
XX
PS   Example 1; Page 27-28; 39pp; French.
XX
CC   The present sequence encodes an outer membrane protein A (OmpA),
CC   designated P40. Enterobacterium OmpA proteins, associated with an
CC   immunogenic peptide from respiratory syncytial virus (RSV), are used to
CC   prepare a nasal composition that induces a protective response, against
CC   RSV infection in the upper and lower (lung) respiratory tract. OmpA
CC   potentiates the immune response to some immunogenic peptides, eliminating
CC   the need for adjuvants. The method is useful for producing vaccines for
CC   prevention or treatment of RSV infections
XX
SQ   Sequence 1032 BP; 225 A; 283 C; 288 G; 236 T; 0 U; 0 Other;

Query Match   99.7%;   Score 1032; DB 4; Length 1032;
Best Local Similarity 100.0%; Pred. No. 2.4e-282;
Matches 1032   Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
```

```
Qy     61  AAACTGGGT  ...
```

```
Qy     61  AAACTGGGTTGGTCCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGACAACAAC  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AAACTGGGTTGGTCCAGTATCACGACACCGGTTCTACGGTAACGGTTTCCAGACAACAAC  120

Qy    121  AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGGCGTTCGGTCGTGGTCGTTACCAGTTAAC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGGCGTTCGGTCGTGGTCGTTACCAGTTAAC  180

Qy    181  CCGTACCTCGGTTCGAAATGGTTATGACGTGGCTGGGCCGTATGCATATAAAGGCAGC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCGTACCTCGGTTCGAAATGGTTATGACGTGGCTGGGCCGTATGCATATAAAGGCAGC  240

Qy    241  GTTGACAACGGTGCTTCAAAGCTCAGGCGGTTCAGCTGACCGCTAAACTGGGTTACCG  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GTTGACAACGGTGCTTCAAAGCTCAGGCGGTTCAGCTGACCGCTAAACTGGGTTACCG  300

Qy    301  ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCATGGTTTGGCCGCGTCGACTCC  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCCATGGTTTGGCCGCGTCGACTCC  360

Qy    361  AAAGGCAACTACGCTTCTACCGGCGTTCCGTAGCGAACACGACACTGGCGTTTCCCCA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AAAGGCAACTACGCTTCTACCGGCGTTCCGTAGCGAACACGACACTGGCGTTTCCCCA  420

Qy    421  GTATTTGCTGCGGCGTAGAGTGGGCTGTTACTCGTGACAATCGCTACCCGTCTGGAATAC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GTATTTGCTGCGGCGTAGAGTGGGCTGTTACTCGTGACAATCGCTACCCGTCTGGAATAC  480

Qy    481  CAGTGGGTTAACAACATCGGCGACGGGCACTGTGGGTACCGTCCTGATAACGGCATG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CAGTGGGTTAACAACATCGGCGACGGGCACTGTGGGTACCGTCCTGATAACGGCATG  540

Qy    541  CTGAGCCTGGGCGTTTCCTACCGGCGTTTCGGTCAGGAAGATGCTGACCGGTTGTTGCTCCG  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CTGAGCCTGGGCGTTTCCTACCGGCGTTTCGGTCAGGAAGATGCTGACCGGTTGTTGCTCCG  600

Qy    601  GCTCCGGCTCCGGCTCCGGCTCCGGAAGTCTACCCTGAAACCGGAAGTCAGCAGGCTCTG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  GCTCCGGCTCCGGCTCCGGCTCCGGAAGTCTACCCTGAAACCGGAAGTCAGCAGGCTCTG  660

Qy    661  TTCAACTTCAACAAGTACCCTGAAACCGGAAGTCAGCAGGCTCGGAACCTGATCAGCTGTAC  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  TTCAACTTCAACAAGTACCCTGAAACCGGAAGTCAGCAGGCTCGGAACCTGATCAGCTGTAC  720

Qy    721  ACTCAGCTGAGCAACATGGATCCGAAAGAGACGGTTCCGCTGTTCTGGGCTACACCGAC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  ACTCAGCTGAGCAACATGGATCCGAAAGAGACGGTTCCGCTGTTCTGGGCTACACCGAC  780

Qy    781  CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840

Qy    841  GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAATCTCCGCGTCGCGGCGCATGGTGAA  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAATCTCCGCGTCGCGGCGCATGGTGAA  900

Qy    901  TCCAACCCGGTTACTGCCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATCGAT  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TCCAACCCGGTTACTGCCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATCGAT  960

Qy    961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020

Qy   1021  CAGCCGCGGGGT  1032
           ||||||||||||
Db   1021  CAGCCGCGGGGT  1032
```

RESULT 9
AAV13868
ID   AAV13868 standard; DNA; 1035 BP.
XX

```
Qy      1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
```

```
AC    AAV13868;
XX
DT    16-SEP-1998 (first entry)
XX
DE    Gene coding for LP40, a variant of OmpA protein P40 from K. pneumoniae.
XX
KW    Outer membrane protein; OmpA; LP40; immunocomplex; oligosaccharide;
KW    polysaccharide; vaccine; Salmonella; ss.
XX
OS    Synthetic.
OS    Klebsiella pneumoniae.
XX
FH    Key            Location/Qualifiers
FT    CDS            1..1035
FT                   /*tag= a
FT                   /product= "LP40"
XX
PN    WO9741888-A1.
XX
PD    13-NOV-1997.
XX
PF    06-MAY-1997;  97WO-FR000800.
XX
PR    07-MAY-1996;  96FR-00005692.
XX
PA    (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI    Binz H,  Haeuw J,  Svenson S;
XX
DR    WPI; 1997-558694/51.
DR    P-PSDB; AAR93797.
XX
PT    Immunogenic complex for use in anti-bacterial vaccine - comprises
PT    bacterial oligo- or poly:saccharide coupled to a Gram-negative bacterial
PT    outer membrane protein or a Streptococcal HSA binding protein.
XX
PS    Disclosure; Page 36-38; 63pp; French.
XX
CC    The patent discloses a new immunogenic complex which consists of (1) an
CC    oligo- or polysaccharide found naturally on bacteria, coupled to (2) a
CC    carrier protein chosen from (a) the human serum albumin binding protein
CC    of Streptococcus, (b) Gram-negative bacterial outer membrane proteins
CC    (Omp), or (c) fragments of these proteins. The immunogenic complex is
CC    useful in a vaccines to protect animals to protect animals against infection by Salmonella,
CC    especially those belonging to antigenic specificity group O:9, including
CC    S. enteritidis, S. panama and S. dublin. A vaccine prepared using an
CC    oligosaccharide from S. enteritidis can be used to provide protection
CC    against septicaemia caused by S. typhi and against typhoid fever, as well
CC    as to protect humans and animals from toxic infections and zoonosis
CC    caused by Salmonella of the same serogroup. The carrier proteins enhance
CC    the immunogenicity of the oligo- or polysaccharide antigens. Inclusion of
CC    additional Salmonella capsule antigens, such as the Vi antigen, increases
CC    the vaccine's efficacy against encapsulated bacteria. The present
CC    sequence represents a gene coding for LP40, a preferred example of a
CC    carrier protein which can be used in the immunocomplex. LP40 is a variant
CC    of the P40 protein from the outer membrane protein of Klebsiella
CC    pneumoniae I-145. The gene can be used for recombinant expression of LP40
CC    in E. coli
XX
SQ    Sequence 1035 BP; 228 A; 284 C; 287 G; 236 T; 0 U; 0 Other;

Query Match       99.5%;   Score 1030.2;   DB 2;   Pred. No. 8e-282;   Length 1035;
Best Local Similarity  99.7%;
Matches  1032;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps  0;

Qy        1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy       61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAGACAAC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAGACAAC 120
```

```
Qy      121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGGTTCGGTGGTTACCAGGTTAAC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGGTTCGGTGGTTACCAGGTTAAC 180

Qy      181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy      241 GTTGACAACGGTGCTTCAAAGCTTCAGCTGCGTTCAGCTACCGCTCAAACTGGGTTACCCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GTTGACAACGGTGCTTCAAAGCTTCAGCTGCGTTCAGCTACCGCTCAAACTGGGTTACCCG 300

Qy      301 ATCACTGACGATCTGGACATCTCACACCGTCTCGGCGGCGGCATGGTTTGGCGCGCTGACTCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATCACTGACGATCTGGACATCTCACACCGTCTCGGCGGCGGCATGGTTTGGCGCGCTGACTCC 360

Qy      361 AAAGGCAACTACGGTTCTACCGGCGGTTTCCCGTAGCGAACACGACACTGCGTTTCCCCA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AAAGGCAACTACGGTTCTACCGGCGGTTTCCCGTAGCGAACACGACACTGCGTTTCCCCA 420

Qy      421 GTATTTGCTGGCGGCGTAGAGTGGCGTTACTCGTGACATCGTCTGACATCGCTCTGAATAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GTATTTGCTGGCGGCGTAGAGTGGCGTTACTCGTGACATCGTCTGACATCGCTCTGAATAC 480

Qy      481 CAGTGGGTTAACACATCCGGCGACGCGGGCACTGTACCCGTCCTGATAACGGCATG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CAGTGGGTTAACACATCCGGCGACGCGGGCACTGTACCCGTCCTGATAACGGCATG 540

Qy      541 CTGAGCCTGGGCGTTCTACCGGTTCGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 CTGAGCCTGGGCGGTTCCTACCGGTTCGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600

Qy      601 GCTCCCGGCTCCGGCTCCGGAAGTGCTACCAACCGGAAGTCAGCAGAGTCTCGGATCAGCTGTAC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 GCTCCCGGCTCCGGCTCCGGAAGTGCTACCAACCGGAAGTCAGCAGAGTCTCGGATCAGCTGTAC 660

Qy      661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGTCGTGGATCAGCTGTAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGTCGTGGATCAGCTGTAC 720

Qy      721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGGTTCCGGTTGTTGTTCTCGGCTACACCGAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGGTTCCGGTTGTTGTTCTCGGCTACACCGAC 780

Qy      781 CGGATCCGGTTCCGAAGCTTACAACCAGCAGCAGCAGCAGTCTGTCTGACGAAAAGCTGCTCAGTCCGTTGTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CGGATCCGGTTCCGAAGCTTACAACCAGCAGCAGCAGCAGTCTGTCTGACGAAAAGCTGCTCAGTCCGTTGTT 840

Qy      841 GACTACCTGGTTGCTAAAGGCATCCGCGGCTGGCAAAATCCGCGCTCGCGGCATGGGTGAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GACTACCTGGTTGCTAAAGGCATCCGCGGCTGGCAAAATCCGCGCTCGCGGCATGGGTGAA 900

Qy      901 TCCAACCGGTTACTGGCAACACACTGTGACAACGTGAAAGCTGAAAGCTGCGCGCCTGCCCTGATCGAT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TCCAACCGGTTACTGGCAACACACTGTGACAACGTGAAAGCTGAAAGCTGCGCGCCTGCCCTGATCGAT 960

Qy      961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGATCGAAGATCGAAGAGTTGTAACT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGATCGAAGATCGAAGAGTTGTAACT 1020

Qy     1021 CAGCCGGCGGGTAA 1035
            |||||  ||  |||
Db     1021 CAGCCGCAGGGCTAA 1035
```

```
RESULT 10
AAF80153
ID   AAF80153 standard; DNA; 1356 BP.
XX
AC   AAF80153;
XX
DT   11-JUN-2001  (first entry)
XX
```

```
Db   301 ATCACTGACGATCTGACATCTACACCCGTCTGGGCGGCTATGGTTTGGCGCCTGACTCC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCGTAGCGAACGACACTGGCGTTTCCCCA 420
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCGTAGCGAACACGACACTGGCGTTTCCCCA 420
Qy   421 GTATTTGCTGCGGCGGCTAGAGTGGGCTGTTACTGTGACATCGCTCTGACTACCCGTCTGGAATAC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGCGGCGGCTAGAGTGGGCTGTTACTGTGACATCGCTCTGACTACCCGTCTGGAATAC 480
Qy   481 CAGTGGTTAACAACATCGGCGACCGCGGCACCGTGGGTCACCGTCCTGATAACGGCATG 540
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGCGACCGCGGCACCGTGGGTCACCGTCCTGATAACGGCATG 540
Qy   541 CTGAGCCTGGCGCGTTCCTACCGCGTTCGGTCAGGAGAGATGTGCACCGGTTGTTGCTCCG 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGCGCGTTCCTACCGCGTTCGGTCAGGAGAGATGTGCACCGGTTGTTGCTCCG 600
Qy   601 GCTCCGCGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGCGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
Qy   661 TTCAACTTCAACAAAGCTACCTACCGGAAGGCTCAGCAGGCTCTCGGATCAGCTGTAC 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGCTACCTACCGGAAGGCTCAGCAGGCTCTCGGATCAGCTGTAC 720
Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTCGTTGTTCTGGGCTACACCGAC 780
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTCGTTGTTCTGGGCTACACCGAC 780
Qy   781 CGCATCGGTTCCGAAGTTCACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db   781 CGCATCGGTTCCGAAGTTCACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTCGTT 840
Qy   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAAATCTCCGCTTCGCGCGGATGGGTGAA 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAAATCTCCGCTTCGCGCGGATGGGTGAA 900
Qy   901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTGCGCGGCTGCCCGATCGAT 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTGCGCGGCTGCCCGATCGAT 960
Qy   961 TGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
Qy  1021 CAGCCCGCCGGG 1031
             |||||| ||||
Db  1021 CAGCCCCAGGG 1031
```

```
DE   Nucleotide sequence of a fusion protein of P40 and RSV antigen.
XX
KW   Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;
KW   RSV; RSV infection; lung; respiratory tract; vaccine; ss.
XX
OS   Synthetic.
OS   Klebsiella pneumoniae.
OS   Respiratory syncytial virus.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1356
FT                  /tag= a
FT                  /product= "fusion protein of P40 and RSV antigen"
FT                  /note= "no termination codon"
XX
PN   WO20012103-A1.
XX
PD   29-MAR-2001.
XX
PF   22-SEP-2000; 2000WO-FR002626.
XX
PR   23-SEP-1999; 99FR-00011888.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Corvaiea N, Goestch L;
XX
DR   WPI; 2001-257929/26.
DR   P-PSDB; AAB67771.
XX
PT   Vaccine against respiratory syncytial virus, comprises enterobacterial
PT   outer membrane protein and viral immunogen, provides protective response
PT   throughout the respiratory tract.
XX
PS   Example 2; Page 29-31; 39pp; French.
XX
CC   The present sequence encodes a fusion protein comprising a Klebsiella
CC   pneumoniae outer membrane protein A (OmpA) designated P40 and a
CC   respiratory syncytial virus (RSV) antigen. Enterobacterium OmpA proteins,
CC   associated with an immunogenic peptide from RSV are used to prepare a
CC   nasal composition that induces a protective response, against RSV
CC   infection in the upper and lower (lung) respiratory tract. OmpA
CC   potentiates the immune response to some immunogenic peptides, eliminating
CC   the need for adjuvants. The method is useful for producing vaccines for
CC   prevention or treatment of RSV infections
XX
SQ   Sequence 1356 BP; 350 A; 397 C; 343 G; 266 T; 0 U; 0 Other;

Query Match             99.1%;  Score 1026.2;  DB 4;  Length 1356;
Best Local Similarity   99.7%;  Pred. No. 1.2e-280;
Matches 1028  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
Qy    61 AAACTGGGTTGGTGTCCCAGTATCACGGACACACCGGTTCTACCGGTAACGGTTCCAGAACAAC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAACTGGGTTGGTGTCCCAGTATCACGGACACACCGGTTCTACCGGTAACGGTTCCAGAACAAC 120
Qy   121 AACCGGTTCCGACCCGTAACGATCAGCGTGGTGCTGGTGCCGGTTCGGTGGTTAAC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCGGTTCCGACCCGTAACGATCAGCGTGGTGCTGGTGCCGGTTCGGTGGTTAAC 180
Qy   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCCCGTATGGCATATAAAGGCAGC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCCCGTATGGCATATAAAGGCAGC 240
Qy   241 GTTGACAACGGTGCTTTCAAAGCTGCTTCAGCTGACCGCTAAACTGGGTTACCCG 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTTCAAAGCTGCTTCAGCTGACCGCTAAACTGGGTTACCCG 300
Qy   301 ATCACTGACGATCTGACATCTACACCCGTCTGGGCGGCTATGGTTTGGCGCCTGACTCC 360
```

```
RESULT 11
AAZ30477
ID   AAZ30477 standard; DNA; 1032 BP.
XX
AC   AAZ30477;
XX
DT   18-JAN-2000  (first entry)
XX
DE   K.pneumoniae OmpA gene.
XX
KW   Outer membrane protein; ompA; enterobacterium; Klebsiella pneumoniae;
KW   immunity; mammal; antigen; hapten; infection; human; bovine;
KW   respiratory syncytial virus; protective response; ss.
XX
OS   Klebsiella pneumoniae.
XX
PN   FR2776521-A1.
XX
PD   01-OCT-1999.
XX
PF   27-MAR-1998;  98FR-00003814.
```

```
XX
PR   27-MAR-1998;  98FR-00003814.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Andreoni C, Rauly I, Nguyen TN, Haeuw JF, Baussant T;
XX
DR   WPI; 1999-583089/50.
DR   P-PSDB; AAY44077.
XX
PT   Immunogenic composition containing bacterial outer membranr protein
PT   conjugated or fused to antigen or hapten, for nasal administration, to
PT   protect against respiratory pathogens.
XX
PS   Disclosure; Page 17-18; 64pp; French.
XX
CC   The invention relates to the use of a fragment of a bacterial membrane
CC   protein, especially the outer membrane protein from an enterobacterium,
CC   e.g. the OmpA protein from Klebsiella pneumoniae, in a composition for
CC   nasal administration to improve immunity, in mammals, against an antigen
CC   or hapten. The antigen or hapten is derived from bacteria and viruses
CC   that cause respiratory infections e.g. the human or bovine respiratory
CC   syncytial virus. This sequence corresponds to the gene encoding the OmpA
CC   protein from Klebsiella. The antigens are shown in AAY44078-Y44149
CC   (AAZ30478-230538 for coding sequences). The use of a membrane protein,
CC   from a species other than that from which the antigen is derived, induces
CC   a protective response against the antigen, even without an adjuvant,
CC   since most adults will already be sensitized against the membrane
CC   protein, although the membrane protein-antigen product will induce an
CC   anti-membrane protein response even in subjects who are not pre-
CC   sensitized
XX
SQ   Sequence 1032 BP; 226 A; 285 C; 285 G; 236 T; 0 U; 0 Other;


Query Match    98.9%;   Score 1024;   DB 2;   Length 1032;
Best Local Similarity   99.5%;   Pred. No. 4.6e-280;
Matches 1027;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

Qy    1 ATGAAAGCAATTTCGTACTGAATGCGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTCGTACTGAATGCGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCACCGGTTAACGGTAACACCACCGGTTCCAGAACAAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCACCGGTTAACGGTAACACCACCGGTTCCAGAACAAC 120

Qy  121 AACGGTCGACCCGTAACGATCAGCTTGGTGTCCGTTCGGTGGTGTCGGTTACCAGGTTAAC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACGGTCGACCCGTAACGATCAGCTTGGTGTCCGTTCGGTGGTGTCGGTTACCAGGTTAAC 180

Qy  181 CCGTACCTCGGTTGGAAATGGGTTATGACTGGCTGGCCGTGGCCGTATGGCGGCATATAAAGGCAGC 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGTACCTCGGTTGGAAATGGGTTATGACTGGCTGGCCGTGGCCGTATGGCGGCATATAAAGGCAGC 240

Qy  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCCGCTAAACTGGGTTACCCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCCGCTAAACTGGGTTACCCG 300

Qy  301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGCGGCCGGCATGGTTTGGCGCGCTGACTCC 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGCGGCCGGCATGGTTTGGCGCGCTGACTCC 360

Qy  361 AAAGGCAACTACGCTTCTACCCGGCGTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGGCAACTACGCTTCTACCCGGCGTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420

Qy  421 GTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy  481 CAGTGGGTTAACAACATCGGGCACGCGGCGACGCGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CAGTGGGTTAACAACATCGGGCACGCGGCGACGCGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540

Qy  541 CTGAGCCTGGGCGTTCCTACCCGTTCGGTCAGAAGATGCTGCACCGGTTGTTGTTGCTCCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTGAGCCTGGGCGTTCCTACCCGTTCGGTCAGGAGAGATGCTGCACCGGTTGTTGTTGCTCCG 600

Qy  601 GCTCCGGCTCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCCGGCTCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy  661 TTCAACTTCAACAAAGTACACCCTGAAACCGAAGTCAGCAGCTCTGGATCAGCTGTAC 720
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TTCAACTTCAACAAAGTACACCCTGAAACCGAAGTCAGCAGCTCTGGATCAGCTGTAC 720

Qy  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCGGGCTACACCGAC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCGGGCTACACCGAC 780

Qy  781 CGCATCCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CGCATCCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy  841 GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACTACCTGGTTGCTAAAGGCATCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900

Qy  901 TCCAACCCGGTTACTGGCAACACTGTGACAACGTGAAAGCTCGGCCTGCCGATCGAT 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCAACCCGGTTACTGGCAACACTGTGACAACGTGAAAGCTCGGCCTGCCGATCGAT 960

Qy  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy 1021 CAGCCGGCGGGT 1032
        |||||    |||
Db 1021 CAGCCTCAGGCT 1032


RESULT 12
AAH78461
ID   AAH78461 standard; DNA; 1032 BP.
XX
AC   AAH78461;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Nucleotide sequence of a Klebsiella protein.
XX
KW   zwitterionic detergent; immune response; vaccine; IgA production;
KW   IgG response; tumour; viral infection; bacterial infection;
KW   parasitic infection; ss.
XX
OS   Klebsiella pneumoniae.
XX
PN   FR2805163-A1.
XX
PD   24-AUG-2001.
XX
PF   21-FEB-2000; 2000FR-00002104.
XX
PR   21-FEB-2000; 2000FR-00002104.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Goestch L,  Corvaia N,  Beck A,  Haeuw JF;
XX
DR   WPI; 2001-591759/67.
DR   P-PSDB; AAG67743.
XX
PT   Mucosally administered vaccines containing zwitterionic detergents to
PT   induce or improve immune response towards antigen or hapten, especially
PT   used in antibacterial, antiviral, antiparasitic or antitumor vaccines.
XX
```

PS   Disclosure; Page 20-21; 26pp; French.
XX
CC   The present sequence encodes a Klebsiella protein. The protein is
CC   administered to test the adjuvant effect of zwitterionic detergents. The
CC   specification describes the use of zwitterionic detergents in the
CC   preparation of a mucosally administered pharmaceutical composition for
CC   inducing or improving the immune response of a mammal towards an antigen
CC   or hapten. In presence of zwitterionic detergents, vaccines are effective
CC   on mucosal (especially nasal) administration, which has the general
CC   advantages of inducing specific IgA production directly at the site of
CC   infection, stimulating a systemic IgG-type response (creating a secondary
CC   barrier against infection) and being simpler to carry out than
CC   administration by injection. The pharmaceutical composition is
CC   specifically a vaccine for the treatment or prophylaxis of tumours or
CC   viral, bacterial or parasitic infections
XX
SQ   Sequence 1032 BP; 226 A; 285 C; 285 G; 236 T; 0 U; 0 Other;

Query Match   98.9%; Score 1024; DB 4; Length 1032;
Best Local Similarity   99.5%; Pred. No. 4.6e-280;
Matches 1027; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

Qy   1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGTTTCTACGGTAACGGTTTCCAGAACAAC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGCGTTGGTTACCAGGTTAAC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGCGTTGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCCCGTATGCATATAAAGGCAGC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGCCCGTATGCATATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTCAAAGCTCAGGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAAGGAACAACGGTGCTTCTACCGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy   301 ATCACTGACGATCTGACATCTGGACACTACACCCGTCTGGGCGGGCATGGTTTGGCGCGCTGACTCC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGACATCTACACCCGTCTGGGCGGGCATGGTTTGGCGCGCTGACTCC 360

Qy   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA 420

Qy   421 GTATTTGCTCGGCGGCTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTCGGCGGCTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy   481 CAGTGGGTTAACAACATCGGGGGACGCGGGACACTGGGTACCCGTCCTGATAACGGCATG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGGGGACGCGGGACACTGGGTACCCGTCCTGATAACGGCATG 540

Qy   541 CTGAGCCTGGGCGTTCCTACCGGCGTTCGGTCAGGAAGATGCTGCACCCGGTTGTTGCTCCG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGTTCCTACCGGCGTTCGGTCAGGAAGATGCTGCACCCGGTTGTTGCTCCG 600

Qy   601 GCTCCGGCTCCGGCGGAAGTGGCTACCAAAGCACTTCACCCTGAGTCTGACGTTCTG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGGCTCCGGCGGAAGTGGCTACCAAAGCACTTCACCCTGAGTCTGACGTTCTG 660

Qy   661 TTCAACTTCAACAAAGTCACCCTGAAACCGGAAGGCTCCGGATCAGCAGCTGTAC 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGTCACCCTGAAACCGGAAGGCTCGGATCAGCAGCTGTAC 720

Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACCGGTTCCGCTGTTGTTCGGGCTACACCGAC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACCGGTTCCGCTGTTGTTCGGGCTACACCGAC 780

Qy   781 CGCATCGGTTCCGAAGCTTCACAACCAGCAGCAGCTGTCTGACAACGTGCTCAGTCCGTTGTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATCGGTTCCGAAGCTTCACAACCAGCAGCAGCTGTCTGACAACGTGCTCAGTCCGTCGTT 840

Qy   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGCCAAAATCTCCGCTCGGCGGCATGGGTGAA 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGCCAAAATCTCCGCTCGGCGGCATGGGTGAA 900

Qy   901 TCCAACCCGGTACTGGCAACACCTGTGACAACGTGAAAGTTCGCGCTGCCCTGATCGAT 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTACTGGCAACACCTGTGACAACGTGAAAGTTCGCGCTGCCCTGATCGAT 960

Qy   961 TGCCTGGCTCCCGGATCGTCTGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCTGGCTCCCGGATCGTCTGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGCGGGT 1032
            |||||||||||
Db   1021 CAGCCTCAGGCT 1032

RESULT 13
AAF90077
ID   AAF90077 standard; DNA; 1032 BP.
XX
AC   AAF90077;
XX
DT   06-AUG-2001 (first entry)
XX
DE   Nucleotide sequence of a recombinant outer membrane protein A (rP40).
XX
KW   Outer membrane protein A; rP40; OMP; enterobacterium; nasal vaccine;
KW   respiratory tract pathogen; ss.
XX
OS   Klebsiella pneumoniae.
XX
PN   FR2801219-A1.
XX
PD   25-MAY-2001.
XX
PF   18-SEP-2000; 2000FR-00011862.
XX
PR   27-MAR-1998; 98FR-00003814.
XX
PA   (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI   Andreoni C, Rauly I, Nguyen T, Haeuw JF, Baussant T;
XX
DR   WPI; 2001-358083/38.
DR   P-PSDB; AAB84122.
XX
PT   Recombinant production of a protein, for particularly use as a carrier
PT   protein in nasal vaccines, comprises renaturation, after extraction, in
PT   the presence of specific detergents.
XX
PS   Example 1; Page 26-27; 48pp; French.
XX
CC   The present sequence encodes a recombinant outer membrane protein A,
CC   designated rP40. The specification describes a method for the recombinant
CC   production of a protein, or its fragments. The method comprises
CC   renaturation the protein, after extraction, in the presence of a solution
CC   containing one of the detergents Zwittergent 3-14, Zwittergent 3-12 and
CC   octylglucopyranoside, where the protein is not beta-interferon. The
CC   protein is especially the outer membrane protein (OMP) from an
CC   enterobacterium, and is useful as a carrier for delivering antigens or
CC   haptens in a nasal vaccine formulation, particularly directed against
CC   pathogens of the respiratory tract, e.g. respiratory syncytial virus
CC   (RSV), (para)influenza, hanta virus, Streptococci, Pneumococci and
CC   Meningococci, especially human or bovine RSV. Most adults show an
CC   antibody response to the Klebsiella pneumoniae P40, and this
CC   presensitisation stimulates the immune response to an antigen/hapten
CC   conjugated to the protein

Db   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTAAAGGCTACAAAGAGAGTTGTAACT 1020

Qy   1021 CAGGCCGGCGGGT 1032
              ||||| |||||
Db   1021 CAGGCCTCAGGCT 1032

RESULT 14
ACH94485
ID   ACH94485 standard; DNA; 1158 BP.
XX
AC   ACH94485;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Klebsiella pneumoniae polynucleotide seqid 280.
XX
KW   Recombinant expression vector; transcription regulatory element;
KW   Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
OS   Klebsiella pneumoniae.
XX
PN   US6610836-B1.
XX
PD   26-AUG-2003.
XX
PF   27-JAN-2000; 2000US-00489039.
XX
PR   29-JAN-1999;   99US-0117747P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton GL, Osborne M;
XX
DR   WPI; 2003-895346/82.
DR   P-PSDB; AB060934.
XX
PT   New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT   preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS   Disclosure; SEQ ID NO 280; 932pp; English.
XX
CC   The invention describes a new isolated nucleic acid encoding a Klebsiella
CC   pneumoniae polypeptide. Also described are: a recombinant expression
CC   vector comprising the nucleic acid, operably linked to a transcription
CC   regulatory element; and a cell comprising the recombinant expression
CC   vector. The nucleic acid is useful for preparing a vaccine composition
CC   against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC   pneumoniae polypeptide of the invention
XX
SQ   Sequence 1158 BP; 259 A; 315 C; 322 G; 262 T; 0 U; 0 Other;

Query Match   97.6%; Score 1009.8; DB 11; Length 1158;
Best Local Similarity   99.3%;   Pred. No. 5.1e-276;
Matches 1014;   Mismatches 7;   Conservative 0;   Indels 0; Gaps 0;

Qy   15 CGTACTGAATGCGGTCCGAAAGATAACACCTGGTATGCGAGGTGTAAACTGGGTGGGTTC 74
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   138 CGTAGCGCAGGCCGCTCCGAAAGATAACACCTGGTATGCGAGGTGTAAACTGGGTGGGTTC 197

Qy   75 CCAGTATCACGACACCGGTTTCTACGGTTAACGGTTTCCAGAACAACAACGGTCCGACCCG 134
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 CCAGTATCACGACACCGGTTTCTACGGTTAACGGTTTCCAGAACAACAACGGTCCGACCCG 257

Qy   135 TAACGATCAGCTTGGTGGTCGTTCGTGGTGGTTAACCCGTTAACCCGGTTAACCTCGGTTT 194
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 TAACGATCAGCTTGGTGGTCGTTCGTGGTGGTTAACCCGTTAACCCGGTTAACCTCGGTTT 317

Qy   195 CGAAATGGGTTATGACTGGCCGTGGCCGTGGCCGTATAAAGGACGTTGACAACGGTGC 254
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 CGAAATGGGTTATGACTGGCCGTGGCCGTGGCCGTATAAAGGACGTTGACAACGGTGC 377

Qy   255 TTTCAAAGCTCAGGGCGTTCAGGTCGATCACTGACGATCT 314
          ||||||||||||||

XX
SQ   Sequence 1032 BP; 226 A; 285 C; 285 G; 236 T; 0 U; 0 Other;

Query Match   98.9%; Score 1024; DB 5; Length 1032;
Best Local Similarity   99.5%;   Pred. No. 4.6e-280;
Matches 1027;   Mismatches 0;   Conservative 0;   Indels 5; Gaps 0;

Qy   1 ATGAAAGCAATTTCGTACTGAATGCGGTCCGAAAGATAACACCTGGTATGCGAGGTGT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTCGTACTGAATGCGGTCCGAAAGATAACACCTGGTATGCGAGGTGT 60

Qy   61 AAACTGGGTGGTCCCCAGTATCACGACACCGGTTTCTACGGTTAACGGTTTCCAGAACAAC 120
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCCAGTATCACGACACCGGTTTCTACGGTTAACGGTTTCCAGAACAAC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCGGTTCGGTGGTGCGGTTACCAGGTTAAC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCGGTTCGGTGGTGCGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTTCGGTTTCGAAATGGGTTATGACTGGCCGTGGCCGTATGGCCGTATAAAGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTTCGGTTTCGAAATGGGTTATGACTGGCCGTGGCCGTATGGCCGTATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCGTAAACTGGGTTACCCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCGTAAACTGGGTTACCCCG 300

Qy   301 ATCACTGACGATCTGACATTCTACACCGCGTCTCGGGCGGCCATGGTTTGGCGCCTGACTCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGACATTCTACACCGCGTCTCGGGCGGCCATGGTTTGGCGCCTGACTCC 360

Qy   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGGCGTTTCCCCA 420

Qy   421 GTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy   481 CAGTGGGTTAACAACATCGGGCACGCGGGCACTGTGGGTACCCGTCCTCGATAACGGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGGCACGCGGGCACTGTGGGTACCCGTCCTCGATAACGGCATG 540

Qy   541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCGTCAGGAGATGCTGTGCACCGGTTGTGCTCCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCGTCAGGAGATGCTGTGCACCGGTTGTGCTCCG 600

Qy   601 GCTCCGGCGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGGCGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy   661 TTCAACTTCAACAAAGCTACCCTGAAACCGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGCTACCCTGAAACCGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTGTTGTTCCGGCTACACCGAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTGTTGTTCCGGCTACACCGAC 780

Qy   781 CGCATCCGGTTCCGAAGCTTACACAACCGAGCAGCTGTCTGAGAAAATCGTGCTCAGTCCGTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATCCGGTTCCGAAGCTTACACAACCGAGCAGCTGTCTGAGAAAATCGTGCTCAGTCCGTT 840

Qy   841 GACTACCTGGTTGCTAAAGGCAATCCCGGCTGGCAAAATCTTCCGCGCTCGCGGCATGGGGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCAATCCCGGCTGGCAAAATCTTCCGCGCTCGCGGCATGGGGAA 900

Qy   901 TCCAACCCGGTTACTGCAACACCTGTGACAACGTGAAAGCTTCGCGCTGCGCCTGATCGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGCAACACCTGTGACAACGTGAAAGCTTCGCGCTGCGCCTGATCGAT 960

Qy   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTAAAGGCTACAAAGAGAGTTGTAACT 1020
          ||||||||||||

PD   26-AUG-2003.
XX
PF   27-JAN-2000; 2000US-00489039.
XX
PR   29-JAN-1999; 99US-0117747P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton GL, Osborne M;
XX
DR   WPI; 2003-895346/82.
DR   P-PSDB; AB061021.
XX
PT   New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT   preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS   Disclosure; SEQ ID NO 367; 932pp; English.
XX
CC   The invention describes a new isolated nucleic acid encoding a Klebsiella
CC   pneumoniae polypeptide. Also described are: a recombinant expression
CC   vector comprising the nucleic acid, operably linked to a transcription
CC   regulatory element; and a cell comprising the recombinant expression
CC   vector. The nucleic acid is useful for preparing a vaccine composition
CC   against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC   pneumoniae polypeptide of the invention
XX
SQ   Sequence 1251 BP; 282 A; 346 C; 348 G; 275 T; 0 U; 0 Other;

Query Match   97.6%; Score 1009.8; DB 11; Length 1251;
Best Local Similarity   99.3%; Pred. No. 5.3e-276;
Matches 1014; Conservative   0; Mismatches   7; Indels   0; Gaps   0;

Qy   15 CGTACTGAATCGGCTCCGAAAGATAACACCTGGTATGCAGTGTGGTAAACTGGGTTGGTC 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 CGTAGCCAGGCCGCGTCCGAAAGATAACACCTGGTATGCAGTGTGGTAAACTGGGTTGGTC 1106

Qy   75 CCAGTATCACGACACCGGTTCTACGTAACGGTTTCGTGGTTAACCAGTTAACAACACGGTTCCGACCCG 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 CCAGTATCACGACACACCGGTTCTACGTAACGGTTTCGTGGTTAACCAGTTAACAACACGGTTCCGACCCG 1046

Qy   135 TAACGATCAGCTTGGTGTCGTGGTTCGGTGGTTCGGTGGTTAACCCGGTTAACCCTCCGTTT 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 TAACGATCAGCTTGGTGTCGTGGTTCGGTGGTTCGGTGGTTAACCCGGTTAACCCTCCGTTT 986

Qy   195 CGAAATGGGTTATGACTGGCTGGGCCGTATGCAGCATATAAAAGGCAGCGTTGACAACGGTGC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   985 CGAAATGGGTTATGACTGGCTGGGCCGTATGCAGCATATAAAAGGCAGCGTTGACAACGGTGC 926

Qy   255 TTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCGGTTACCCTCCGTTT 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   925 TTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCGGTTACCCTGACGATCT 866

Qy   315 GGACATCTACACCCGTCTCGGGCGGCGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGC 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   865 GGACATCTACACCCGTCTCGGGCGGCGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGC 806

Qy   375 TTCTACCGGCGTTCCCGTAGCGAACGACAGACTGGCGTTCCCCAGTATTGCTGGCGG 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   805 TTCTACCGGCGTTCCCGTAGCGAACGACAGACTGGCGTTCCCCAGTATTGCTGGCGG 746

Qy   435 CGTAGAGTGGGCGTGTTACTCGTTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAA 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   745 CGTAGAGTGGGCGTGTTACTCGTTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAA 686

Qy   495 CATCGGCGACGGCGGCACTGTGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGT 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   685 CATCGGCGACGGCGGCACCGTGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGT 626

Qy   555 TTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGCTCCGGCTCCGGC 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   625 TTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCGCTCCGGCTCCGGC 566

Qy   615 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGTTGAACTTCAACAA 674

RESULT 15
ACH94572/c
ID   ACH94572 standard; DNA; 1251 BP.
XX
AC   ACH94572;
XX
DT   29-JUL-2004 (first entry)
XX
DE   Klebsiella pneumoniae polynucleotide seqid 367.
XX
KW   Recombinant expression vector; transcription regulatory element;
KW   Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
OS   Klebsiella pneumoniae.
XX
PN   US6610836-B1.
XX

Db   378 TTTCAAAGCTCAGGGCGGTTCAGCTGACCGCTAAACTGGGTTACCCGGTTACCCTGACGATCT 437

Qy   315 GGACATCTACACCCGTCTCGGGCGGCGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGC 374

Db   438 GGACATCTACACCCGTCTGGGGCGGCGCATGGTTTGGCGCGCTGACTCCAAAGGCAACTACGC 497

Qy   375 TTCTACCGGCGTTCCCGTAGCGAACGACACTGGCGTATTTGCTGGCGG 434

Db   498 TTCTACCGGCGTTCCCGTAGCGAACGACACTGGCGTATTTGCTGGCGCGG 557

Qy   435 CGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATACCAGTGGGTTAACAA 494

Db   558 CGTAGAGTGGGCTGTTACTCGTCGACATCGCTGACATCGCTCGGAATACCAGTGGGTTAACAA 617

Qy   495 CATCGGCGACGGCGGCACCGTGGGGGTACCCGTCCTGATAACGGCATGCTGAGCCTGGGCGT 554

Db   618 CATCGGCGACGGCGGCACCCGTGGGGTACCCGTCCTGATAACGGCATGCCTGGGCCTGGGCGT 677

Qy   555 TTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTGCTCCGCTCCGGCTCCGGC 614

Db   678 TTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTGCTCCGCTCCGGCTCCGGC 737

Qy   615 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGTTGACGTTCGTGTTCTGTTCAACTTCAACAA 674

Db   738 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCTGTTCAACTTCAACAA 797

Qy   675 AGCTACCCTGAAACCGGAAGGTTCAGCGGTTCAGCTGATCAGCTGTACACTGCAGTGAGCAA 734

Db   798 AGCTACCCTGAAACCGGAAGGTTCAGCGGTTCAGCTCAGCAGGCTCTGAGCTGTACACTGCAGTGAGCAA 857

Qy   735 CATGGATCCGCGAAAGACGGTTCCGCTGTGTTCTGGGCTACACCGACCGGCATCGGTTCCGA 794

Db   858 CATGGATCCGCGAAAGACGGTTCCGCTGTGTTCTGGGCTACACCGACCGGCATCGGTTCCGA 917

Qy   795 AGCTTACAACCAGCAGCTGTCGTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 854

Db   918 AGCTTACAACCAGCAGCTGTCGTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 977

Qy   855 TAAAGGCATCCCGGCGTGGCAAAATCCCGCTCCGCTCGCGCGGCATGGTGAATCCAACCCGGTTAC 914

Db   978 TAAAGGCATCCCGGCGTGGCAAAATCCCGCTCCGCTCGCGCGGCATGGTGAATCCAACCCGGTTAC 1037

Qy   915 TGGCAACACCTGTGACAACGTGAAAGCTCGCGCGCTGCCTGATTGCCTGGCTCCGGA 974

Db   1038 TGGCAACACCTGTGACAACGTGAAAGCTCGCGCGCTGCCTGATTGCCTGGCTCCGGA 1097

Qy   975 TCGTCGTGTGTAGAGATCGAAGTAAAGGCTACAAAGAAGTGTAACTCAGCGCGGCGGGTTA 1034

Db   1098 TCGTCGTGTGTAGAGATCGAAGTAAAGGCTACAAAGAAGTGTAACTCAGCGCGGCGGCTTA 1157

Qy   1035 A 1035

Db   1158 A 1158

```
Db  565 TCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTGTTCAACTTCAACAA 506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  675 AGCTACCCTGAAACCGGAAGGTCTCGGAATCAGCTGTACACTCAGCTGAGCAA 734
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  505 AGCTACCCTGAAACCGGAAGGTCTCGGAATCAGCTGTACACTCAGCTGAGCAA 446

Qy  735 CATGGATCCGAAAGACCGGTTCCGCTGTTGTTCTGGGCTACACCGACCGCATCGGTTCCGA 794
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  445 CATGGATCCGAAAGACCGGTTCCCGTGTTGTTCTGGGCTACACCGACCCCATCCGGTTCCGA 386

Qy  795 AGCTTACACACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 854
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  385 AGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTTGACTACCTGGTTGC 326

Qy  855 TAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTAC 914
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  325 TAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAATCCAACCCGGTTAC 266

Qy  915 TGGCAAACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGCTGATTGCCTGGCTCCGGA 974
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  265 TGGCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGCTGATTGCCTGGCTCCGGA 206

Qy  975 TCGTCGTGTGTAGAGATCGAAGTTAAAGGCTACAAAGAGAGTTGTAACTCAGCCGGCGGGTTA 1034
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db  205 TCGTCGTGTGTAGAGATCGAAGTTAAAGGCTACAAAGAGAGTTGTAACTTGTAACTCAGCCGCGGGGCTTA 146

Qy 1035 A 1035
            |
Db  145 A 145
```

Search completed: January 19, 2005, 14:11:58
Job time : 610 secs

This Page Blank (uspto)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2005, 10:19:08 ; Search time 4768 Seconds
                                    (without alignments)
                        10265.277 Million cell updates/sec

Title: US-09-913-772-1
Perfect score: 1035
Sequence: 1 atgaaagcaatttcgtact.........taactcagccgcggcgggtaa 1035

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
          1: gb_ba:*
          2: gb_htg:*
          3: gb_in:*
          4: gb_om:*
          5: gb_ov:*
          6: gb_pat:*
          7: gb_ph:*
          8: gb_pl:*
          9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1035	100.0	1035	6	BD260130	Use of an
2	1035	100.0	1035	6	BD272054	Bacterial
3	1035	100.0	1035	6	BD275000	Immunosti
4	1035	100.0	1035	6	AX033528	Sequence
5	1035	100.0	1035	6	AX033532	Sequence
6	1035	100.0	1035	6	AX034418	Sequence
7	1035	100.0	1035	6	AX035661	Sequence
8	1035	100.0	1035	6	AX035701	Sequence
9	1035	100.0	1035	6	AX191406	Sequence
10	1032	99.7	1035	6	AX101162	Sequence
11	1030.2	99.5	1035	6	KPOUTMEMA	Klebsiell
12	1030.2	99.5	1035	6	A67574	Sequence 3
13	1026.2	99.1	1356	6	AX101164	Sequence
14	1024	98.9	1032	6	BD222264	Use of ac
15	1024	98.9	1032	6	AX016041	Sequence
16	1024	98.9	1032	6	AX230577	Sequence
17	1009.8	97.6	1158	6	AR383551	Sequence
c 18	1009.8	97.6	1251	6	AR383638	Sequence 13
19	1008	97.4	1008	6	A47011	Sequence 13

20	1008	97.4	1008	6	A67572	A67572 Sequence 1
21	1008	97.4	1008	6	AR108918	AR108918 Sequence
22	1008	97.4	1008	6	AR214589	AR214589 Sequence
23	1008	97.4	1008	6	AR287914	AR287914 Sequence
24	1008	97.4	1008	6	AR392815	AR392815 Sequence
25	1008	97.4	1008	6	AX150311	AX150311 Sequence
26	1007	97.3	1007	6	AX138189	AX138189 Sequence
27	1007	97.3	1361	1	A51216	A51216 Sequence 1
28	766.6	74.1	1361	1	X00254	EAOMPA Enterobacte
29	724.4	70.0	730	1	KPNOMPAA	M63355 K.pneumonia
30	724.4	70.0	730	6	AR089252	AR089252 Sequence
31	724.4	70.0	730	1	AR093452	AR093452 Sequence
32	712.2	68.8	1791	1	SDOMPA	V01344 Shigella dy
33	702.6	67.9	13832	1	AF234269	AF234269 Escherich
c 34	697.8	67.4	18269	1	D90734	D90734 Escherichia
c 35	697.8	67.4	110000	1	D90733	D90733 Escherichia
c 36	697.8	67.4	110000	1	U00096_10	Continuation (11 o
c 37	697.8	67.4	301276	1	AE016758	AE016758 Escherich
38	696.2	67.3	1041	1	AY305874	AY305874 Shigella
39	696.2	67.3	1041	1	AY590298	AY590298 Shigella
40	696.2	67.3	2270	1	AX191724	AX191724 Sequence
41	696.2	67.3	2271	1	ECOMPA	V00307 E. coli omp
c 42	694.6	67.1	11641	1	AE005286	AE005286 Escherich
c 43	694.6	67.1	327773	1	AP002554	AP002554 Escherich
44	685.8	66.3	5629	6	A08457	A08457 pHS 164 mRN
c 45	685	66.2	254050	1	AL627269	AL627269 Salmonell

                    ALIGNMENTS

RESULT 1
BD260130
LOCUS   BD260130   1035 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION Use of an enterobacterium protein OmpA for specific targeting
            towards antigen-presenting cells.
ACCESSION   BD260130
VERSION   BD260130.1 GI:33069900
KEYWORDS   JP 2002529428-A/1.
SOURCE   Klebsiella pneumoniae
  ORGANISM Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1 (bases 1 to 1035)
  AUTHORS   Bonnefoy,J.Y., Lecoanet,S., Aubry,J.P., Jeannin,P. and Baussant,T.
  TITLE   Use of an enterobacterium protein OmpA for specific targeting
            towards antigen-presenting cells
  JOURNAL   Patent: JP 2002529428-A 1 10-SEP-2002;
            PIERRE FABRE MEDICAMENT
COMMENT   OS   Klebsiella pneumoniae
            PN   JP 2002529428-A/1
            PD   10-SEP-2002
            PF   08-NOV-1999 JP 2000580661
            PR   06-NOV-1998 FR   98/14007
            PI   JEAN YVES BONNEFOY,SYBILLE LECOANET,JEAN PIERRE AUBRY,PASCALE
            PI   JEANNIN,
            PI   THIERRY BAUSSANT
            PC   A61K38/00,A61K35/76,A61K39/02,A61K39/385,A61K39/39,A61K48/00,
            PC   A61P9/00,
            PC   A61P25/00,A61P29/00,A61P31/00,A61P35/00,A61P37/02,A61P37/04,
            PC   A61P37/06,
            PC   C07K14/26,C07K19/00,C12N5/06,C12N15/09,A61K37/02,C12N15/00, PC
            C12N5/00
            CC   Use of an enterobacterium protein OmpA for specific targeting
            CC   towards
            CC   antigen-presenting cells
            FH   Key   Location/Qualifiers
            FT   exon   (1)..(1032)
            FT   intron   (1033)..(1035)
            FT   CDS   (1)..(1032)
FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Klebsiella pneumoniae"

ORIGIN

/mol_type="genomic DNA"
/db_xref="taxon:573"

Query Match   100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.1e-241;
Matches 1035; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCCAGAACACC 120
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCCAGAACACC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGGTTACCAGGTTAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTGGTTCGAAATGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTGGTTCGAAATGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACTGGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACTGGGTTACCCG 300

Qy   301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCGGCCGGTTGGGCGCGCTGACTCC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCGGCCGGTTGGGCGCGCTGACTCC 360

Qy   361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGGACGACACTGGCGTTTCCCCA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGTTCCCGGACGACACTGGCGTTTCCCCA 420

Qy   421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGTATCGTCACCGGTCTCTGAAATAC 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGTATCGTCACCGGTCTCTGAAATAC 480

Qy   481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTCGTGGGTACCCGTCCTGATAACGGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTCGTGGGTACCCGTCCTGATAACGGCATG 540

Qy   541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600

Qy   601 GCTCCGCGGCTCCGGCTCCGGAAGCTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTCTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGCGGCTCCGGCTCCGGAAGCTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTCTG 660

Qy   661 TTCAACTTCAACAAAGCTCACCCTGAAACCGGAAGCTCGAGGCTCTGAATCAGCTGTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGCTCACCCTGAAACCGGAAGCTCGAGGCTCTGAATCAGCTGTAC 720

Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTTGTTCGGGCTACACCGAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTTGTTCGGGCTACACCGAC 780

Qy   781 CGCATGGCGGTTCCGAAGCTTACAACCAGCACGTGCTCAGCTTCAGTCCGTCCGTTGTT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATGGCGGTTCCGAAGCTTACAACCAGCACGTGCTCAGCTTCAGTCCGTCCGTTGTT 840

Qy   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTCGCAAAATCTCCGCGCCTCGGGCCATGGGTGAA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTCGCAAAATCTCCGCGCCTCGGGCCATGGGTGAA 900

Qy   901 TCCAACCCGGTTACTGCACAACACCTGTGACAACGTGAAAGCTCGCCGCTGCCCTGATCGAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGCACAACACCTGTGACAACGTGAAAGCTCGCCGCTGCCCTGATCGAT 960

Qy   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021 CAGCCGCGGCGGTTAA 1035
          ||||||||||||||||
Db   1021 CAGCCGCGGCGGTTAA 1035

RESULT 2
BD272054
LOCUS   BD272054   1035 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION Bacterial membrane fractions with adjuvant effect.
ACCESSION   BD272054
VERSION   BD272054.1 GI:33081822
KEYWORDS   JP 2002539169-A/1.
SOURCE   Klebsiella pneumoniae
  ORGANISM Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1 (bases 1 to 1035)
  AUTHORS   Libon,C., Corvaia,N., NGguyen,T.N., Beck,A. and Bonnefoy,J.Y.
  TITLE   Bacterial membrane fractions with adjuvant effect
  JOURNAL   Patent: JP 2002539169-A 1 19-NOV-2002;
            PIERRE FABRE MEDICAMENT
COMMENT   OS   Klebsiella pneumoniae
            PN   JP 2002539169-A/1
            PD   19-NOV-2002
            PF   15-MAR-2000 JP 2000604864
            PR   15-MAR-1999 FR   99/03153
            PI   CHRISTINE LIBON, NATHALIE CORVAIA, THIEN NGOC
            N'GUYEN, ALAIN BECK,
            PI   JEAN YVES BONNEFOY
            PC   A61K39/39,A61K9/107,A61K35/74,A61K38/00,A61K39/108,A61K39/155,
            PC   A61K47/48,
            PC   A61P31/00,A61P31/04,A61P31/10,A61P31/12,A61P33/00,A61P35/00,
            PC   A61P37/02,
            PC   C12N1/20//C07K19/00,(C12N1/20,C12R1:20),(C12N1/20,C12R1:22),A61K37/02 CC
            Bacterial membrane fractions with adjuvant efffect FH   Key
            Location/Qualifiers
FT   exon   (1)..(1032)
FT   intron   (1033)..(1035)
FT   CDS   (1)..(1032).
FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:573"
ORIGIN

Query Match   100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.1e-241;
Matches 1035; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCCAGAACAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCCAGAACAAC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGGTTACCAGGTTAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGGTCGTGCGTTCGGTGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTCGGTTTGCTAAAGGCCATCGGTTTGCTAAAGGCCATATGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTTGCTAAAGGCCATCGGTTTGCTAAAGGCCATATGGCAGC 240

Qy   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACTGGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACTGGGTTACCCG 300

```
PR   15-MAR-1999 FR     99/03154
PI   christine libon,nathalie corvaia,allen beck,jan yves vonnefoy
CC
FH   Key          Location/Qualifiers
FT   exon         (1)..(1032)
FT   intron       (1033)..(1035)
FT   CDS          (1)..(1032)

                  Location/Qualifiers
FEATURES
     source       1..1035
                  /organism="Klebsiella pneumoniae"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:573"

ORIGIN

Query Match          100.0%;  Score 1035;  DB 6;  Length 1035;
Best Local Similarity 100.0%;  Pred. No.1.1e-241;
Matches 1035;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60

Qy    61 AAACTGGGTTGGTCGCCGTTTCACGGTAACGGTTTCCAGAACAAC                120
         ||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAACTGGGTTGGTCGCCGTTTCACGGTAACGGTTTCCAGAACAAC                120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTTGTGTGCTGGTGGTGTGGTGGGCCGCATAAAGGCAGC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTTGTGTGCTGGTGGTGTGGTGGGCCGCATAAAGGCAGC  180

Qy   181 CCGTACCTCGGTTCGAAATGGGTTATGACTGGGCTAAACTGGGTTACCCG          240
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCGGTTCGAAATGGGTTATGACTGGGCTAAACTGGGTTACCCG          240

Qy   241 GTTGACAACGGTGCTTTCAAAGCGTGGGCGTTCAGGCGCTCGGGCGGCATGATCC    300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTTCAAAGCGTGGGCGTTCAGGCGCTCGGGCGGCATGATCC    300

Qy   301 ATCACTGACGATCTGGACATCTACACCCGTCGGGCGTTTCCGGCGTTTCCGGC       360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGGACATCTACACCCGTCGGGCGTTTCCGGCGTTTCCGGC       360

Qy   361 AAAGGCAACTACGCTTCTACCGGGCGTTCCCGGCGTTTCCCA                420
         ||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGGCGTTCCCGGCGTTTCCCA                420

Qy   421 GTATTTGCTGGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGTCTGGAATAC     480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGTCTGGAATAC     480

Qy   481 CAGTGGGTTAACAACAATCGGCGACGCGGGGCACTGTACCCGTCCTGATAACGGCATG   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACAATCGGCGACGCGGGGCACTGTACCCGTCCTGATAACGGCATG   540
```

```
Qy   301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTGGCGCTGACTCC   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTGGCGCTGACTCC   360

Qy   361 AAAGGCAACTACGCTTCTACCGGGCGTTCCCGTAGCGAACGACACTGGCGTTTCCCA   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGGCGTTCCCGTAGCGAACGACACTGGCGTTTCCCA   420

Qy   421 GTATTTGCTGGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGTCTGGAATAC     480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGTCTGGAATAC     480

Qy   481 CAGTGGGTTAACAACAATCGGCGACGCGGGGCACTGTGGCTGTGGGTACCCG       540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACAATCGGCGACGCGGGGCACTGTGGCTGTGGGTACCCG       540

Qy   541 CTGAGCCTGGGCGTTCGGCTCCGGCTCCGGAAGTGCTGCACCGGTTGCTCCG       600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGTTCGGCTCCGGCTCCGGAAGTGCTGCACCGGTTGCTCCG       600

Qy   601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG  660

Qy   661 TTCAACTTCAACAAAGTTACCCTGGCAACCGGAAGGTCAGGCTCTGGATCAGCGTGTAC   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGTTACCCTGGCAACCGGAAGGTCAGGCTCTGGATCAGCGTGTAC   720

Qy   721 ACTCAGCTGAGCAACATGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC   780

Qy   781 CGCATCGGTTCCGAAGCTTACAACCAGCTGCTGAGAAACGTGCTCAGTGCTCAGTCCGGTTGTT  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATCGGTTCCGAAGCTTACAACCAGCTGCTGAGAAACGTGCTCAGTGCTCAGTCCGGTTGTT  840

Qy   841 GACTACCTGGTTGCTAAAGGCATCCCCGGCTGGCAAAATCTCCCGCGGCATGGGTGAA   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCCGGCTGGCAAAATCTCCCGCGGCATGGGTGAA   900

Qy   901 TCCAACCCGGTTACTGACAACACCTGTGACAACGTGAAAGCTCGCGGCTGCCCTGATCGAT  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGACAACACCTGTGACAACGTGAAAGCTCGCGGCTGCCCTGATCGAT  960

Qy   961 TGCCTGGCTCCGGATCGTCGTCGTGAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCTGGCTCCGGATCGTCGTCGTGAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020

Qy  1021 CAGCCGGCGGGTTAA  1035
         |||||||||||||||
Db  1021 CAGCCGGCGGGTTAA  1035
```

```
RESULT 3
BD275000
LOCUS       BD275000         1035 bp    DNA      linear   PAT 17-JUL-2003
DEFINITION  Immunostimulant bacterial membrane fractions in cancer treatment.
ACCESSION   BD275000
VERSION     BD275000.1  GI:33084768
KEYWORDS    JP 2002539170-A/1.
SOURCE      Klebsiella pneumoniae
ORGANISM    Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1  (bases 1 to 1035)
AUTHORS     Libon,C., Corvaia,N., Beck,A. and Vonnefoy,J.Y.
TITLE       Immunostimulant bacterial membrane fractions in cancer treatment
JOURNAL     Patent: JP 2002539170-A 1 19-NOV-2002;
            PIERRE FABRE MEDICAMENT
COMMENT     OS   Klebsiella pneumoniae
            PN   JP 2002539170-A/1
            PD   19-NOV-2002
            PF   15-MAR-2000 JP 2000604865
```

```
Qy    181 CCGTACCTCGGTTTCGAAATGGTTATGACTGGCTGGCCTATGCATATAAAGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        CCGTACCTCGGTTTCGAAATGGTTATGACTGGCTGGCCTATGCATATAAAGGCAGC

Qy    241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGGTTCAGCTGACCGCTAAACTGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGGTTCAGCTGACCGCTAAACTGGTTACCCG 300

Qy    301 ATCACTGACGATCTGACACATCACACCGTCGGGCGCGTCTGGGCGGGCATGGTTGGCGGGCGCTGACTCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATCACTGACGATCTGACACATCACACCGTCGGGCGCGTCTGGGCGGGCATGGTTGGCGGGCGCTGACTCC 360

Qy    361 AAAGGCAACTACGCTTCTACCGGCGTTCCGTAGCGAACACGACACTGGCGTTTCCCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGGCAACTACGCTTCTACCGGCGTTCCGTAGCGAACACGACACTGGCGTTTCCCCA 420

Qy    421 GTATTTGCTGGGCGGCTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATTTGCTGGGCGGCTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGCGCACGCGGCACCGTCCGTGATAACGGCATG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGCGCACGCGGCACCGTCCGTGATAACGGCATG 540

Qy    541 CTGAGCCTGGGCGTTCTCCGGCGTTCTCCTACCGGTTCGGTCAGGAAGATGCTGCACCGGTTGCTGCTCCG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGAGCCTGGGCGTTCTCCGGCGTTCTCCTACCGGTTCGGTCAGGAAGATGCTGCACCGGTTGCTGCTCCG 600

Qy    601 GCTCCCGGCCTCCGGAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGGTCTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCCGGCCTCCGGAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGGTCTG 660

Qy    661 TTCAACTTCAACAAGCTACCCTGAAACCGAAACCGGTTCAGCAGGCTCTGGATCAGCTGTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCAACTTCAACAAGCTACCCTGAAACCGAAACCGGTTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy    721 ACTCAGCTGAGCAACATGGATCCGAAAGACCGTTCCGCTGTTGTTCTGGGCTACACCGAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACTCAGCTGAGCAACATGGATCCGAAAGACCGTTCCGCTGTTGTTCTGGGCTACACCGAC 780

Qy    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy    841 GACTACCTGGTTGCTAAAGGCATCCCGGGCTGGCAAAATCTCCGCTGCAAAATCTCCGCTGCCGTGGTGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTACCTGGTTGCTAAAGGCATCCCGGGCTGGCAAAATCTCCGCTGCAAAATCTCCGCTGCCGTGGTGAA 900

Qy    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCCGCTGCCCTGATCGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCCGCTGCCCTGATCGAT 960

Qy    961 TGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCTGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGGCGGGTTAA 1035
          |||||||||||||||
Db   1021 CAGCCGGCGGGTTAA 1035
```

RESULT 5
AX033532
LOCUS   AX033532   1035 bp   DNA   linear   PAT 21-SEP-2000
DEFINITION   Sequence 1 from Patent WO0048629.
ACCESSION   AX033532
VERSION   AX033532.1 GI:10280280
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM   Klebsiella pneumoniae
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Miconnet,I., Romero,P., Carottini,J.C., Renno,T. and Bonnefoy,J.Y.

```
Qy    841 GACTACCTGGTTGCTAAAGGCATCCCGGGCTGCAAAATCTCCGCTGCCGTGGTGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTACCTGGTTGCTAAAGGCATCCCGGGCTGCAAAATCTCCGCTGCCGTGGTGAA 900

Qy    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCCGCTGCCCTGATCGAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCCGCTGCCCTGATCGAT 960

Qy    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGGCGGGTTAA 1035
          |||||||||||||||
Db   1021 CAGCCGGCGGGTTAA 1035
```

RESULT 4
AX033528
LOCUS   AX033528   1035 bp   DNA   linear   PAT 21-SEP-2000
DEFINITION   Sequence 1 from Patent WO0048628.
ACCESSION   AX033528
VERSION   AX033528.1 GI:10280278
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM   Klebsiella pneumoniae
               Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
               Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Renno,T. and Bonnefoy,J.Y.
  TITLE   Use of an enterobacterium protein ompa associated with an antigen
               for generating an antiviral, antiparasitic or antitumoral cytotoxic
               response
  JOURNAL   Patent: WO 0048628-A 1 24-AUG-2000;
               PF MEDICAMENT (FR) ; RENNO TOUFIC (FR) ; BONNEFOY JEAN YVES (FR)
FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS   1..1035
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC09985.1"
                     /db_xref="GI:10280279"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGNSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAFGGYGVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSREHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWNNIGDAGTVGTRPDNGMLSLGVSYRPGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN

Query Match   100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity 100.0%;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy   121 AACGGTTCCGACCCGTAACGACATCAGCTCAGTTGGTCGTCGTCGGTCGGTTCGGTTACCAGGTTAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTTCCGACCCGTAACGACATCAGCTCAGTTGGTCGTCGTCGGTCGGTTCGGTTACCAGGTTAAC 180

Qy   181 CCGTACCTCGGTTTCGAAATGGTTATGACTGGCTGGCCTATGCATATAAAGGCAGC 240
```

```
TITLE       Use of an ompa enterobacterium protein associated
JOURNAL     Patent: WO 0048629-A 1 24-AUG-2000;
            MICONNET ISABELLE (CH) ; ROMERO PEDRO (CH) ; CAROTTINI JEAN CHARLES
            (CH) ; RENNO TOUFIC (FR) ; BONNEFOY JEAN YVES
            (FR)
            PF MEDICAMENT (FR)
FEATURES             Location/Qualifiers
     source          1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS             1..1035
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC09986.1"
                     /db_xref="GI:10280281"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAFGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLRGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLIVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN

Query Match             100.0%;  Score 1035;  DB 6;  Length 1035;
Best Local Similarity   100.0%;  Pred. No. 1.1e-241;
Matches 1035;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 ATGAAAGCAATTTTCGTACTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGAAAGCAATTTTCGTACTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy        61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy       121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTTACCAGGTTAAC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTTACCAGGTTAAC 180

Qy       181 CCGTACCCGGTTTCGAAATGGGTTATGACTGCTGGGCCGTATGGCATATAAAGGCAGC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCGTACCCGGTTTCGAAATGGGTTATGACTGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy       241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCCTAAACTGGGTTACCCG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCCTAAACTGGGTTACCCG 300

Qy       301 ATCACTGACGATCTGGACATCTACACCCGTCTGCGCGGCCGCATGGTTTGGCGCGCTGACTCC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ATCACTGACGATCTGGACATCTACACCCGTCTGCGCGGCCGCATGGTTTGGCGCGCTGACTCC 360

Qy       361 AAAGGCAACTACGGCTTCTACCGGCGTTCCCGGTAGCGAACACGACACTGGCGTTTCCCCA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AAAGGCAACTACGGCTTCTACCGGCGTTCCCGGTAGCGAACACGACACTGGCGTTTCCCCA 420

Qy       421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCCACCGTCTGGAATAC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCCACCGTCTGGAATAC 480

Qy       481 CAGTGGGTTAACAACAATCGGCGACGCGGGCACTGTGGTACCCGTCCTGATAACGGCATG 540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CAGTGGGTTAACAACAATCGGCGACGCGGGCACTGTGGTACCCGTCCTGATAACGGCATG 540

Qy       541 CTGAGCGGCCTGGGCGTTCCTCGGTACCGGCTCCGGAAGATGCTGCACCGGGTTGCTCCG 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CTGAGCGGCCTGGGCGTTCCTCGGTACCGGCTCCGGAAGATGCTGCACCGGGTTGCTCCG 600

Qy       601 GCTCCCGGCTCCGCCTCCGGAAGTGGCTACCAAGCTGAAGTCTGACGTTCTG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GCTCCCGGCTCCGCCTCCGGAAGTGGCTACCAAGCTGAAGTCTGACGTTCTG 660

Qy       661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy       721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC 780

Qy       781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy       841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900

Qy       901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATCGAT 960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCTGCCCTGATCGAT 960

Qy       961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy      1021 CAGCCGGCGGGGTTAA 1035
             ||||||||||||||||
Db      1021 CAGCCGGCGGGGTTAA 1035


RESULT 6
AX034418
LOCUS       AX034418                1035 bp    DNA     linear   PAT 22-SEP-2000
DEFINITION  Sequence 1 from Patent FR2789902.
ACCESSION   AX034418
VERSION     AX034418.1  GI:10303128
KEYWORDS    .
SOURCE      Klebsiella pneumoniae
  ORGANISM  Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella;
REFERENCE   1
  AUTHORS   Haeuw,J.F., Beck,A., Goetsch,L., Bonnefoy,J.Y. and Corvaia,N.
  JOURNAL   Patent: FR 2789902-A 1 25-AUG-2000;
            PF MEDICAMENT.(FR)
FEATURES             Location/Qualifiers
     source          1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS             1..1035
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC10120.1"
                     /db_xref="GI:10303129"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAFGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLRGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLIVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN

Query Match             100.0%;  Score 1035;  DB 6;  Length 1035;
Best Local Similarity   100.0%;  Pred. No. 1.1e-241;
Matches 1035;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         1 ATGAAAGCAATTTTCGTACTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGAAAGCAATTTTCGTACTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy        61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
```

```
VERSION      AX035661.1  GI:11191257
KEYWORDS     .
SOURCE       Klebsiella pneumoniae
  ORGANISM   Klebsiella pneumoniae
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Klebsiella.
REFERENCE    1
  AUTHORS    Beck,A., Libon,C., Bonnefoy,J.Y. and Corvaia,N.
  TITLE      Immunostimulant bacterial membrane fractions in cancer treatment
  JOURNAL    Patent: WO 0054790-A 1 21-SEP-2000;
             PF MEDICAMENT (FR) ; BECK ALAIN (FR) ; LIBON CHRISTINE (FR) ;
             BONNEFOY JEAN YVES (FR) ; CORVAIA NATHALIE (FR)
FEATURES             Location/Qualifiers
     source          1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS             1..1035
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC16264.1"
                     /db_xref="GI:11191258"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAFGGYGVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQOALDQLYTOLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN

Query Match             100.0%;  Score 1035;  DB 6;  Length 1035;
Best Local Similarity   100.0%;
Matches 1035;           Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;  Pred. No. 1.1e-241;

Qy          1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy         61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy        121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCGTGGTGCGTTCGGTGCGTTACCAGTTAAC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCGTGGTGCGTTCGGTGCGTTACCAGTTAAC 180

Qy        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGTTATGACTGGTTATGACTGGTTACCCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGTTATGACTGGTTATGACTGGTTACCCG 240

Qy        241 GTTGACAACGGTGCTTCAAAGCTGCTTCAGGGCGTTCAGCTGACCGGTCAAACTGGGTTACCCG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTGACAACGGTGCTTCAAAGCTGCTTCAGGGCGTTCAGCTGACCGGTCAAACTGGGTTACCCG 300

Qy        301 ATCACTGACGATCTGGACATCTGACATCTCACCGGTTTCTACCGGTAACGGTTTCCACCGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCACTGACGATCTGGACATCTGACATCTCACCGGTTTCTACCGGTAACGGTTTCCACCGAC 360

Qy        361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGGAGCGAACAGCACTGGGCGTTTCCCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGGAGCGAACAGCACTGGGCGTTTCCCCA 420

Qy        421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC 480

Qy        481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGGTGGGTACCACCGTCCTGATAACGGCATG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGGTGGGTACCACCGTCCTGATAACGGCATG 540

Qy        541 CTGAGCCGTGGGCGTTTCCTACCGGTTCGGTCAGGAAGATGTGCACCGGTTGTTGCTCCG 600
```

(left column, rotated)

```
Db         61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy        121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCGTGGTGCGTTCGGTGCGTTACCAGTTAAC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCGTGGTGCGTTCGGTGCGTTACCAGTTAAC 180

Qy        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGTTATGACTGGTTATGACTGGTTACCCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGTTATGACTGGTTATGACTGGTTACCCG 240

Qy        241 GTTGACAACGGTGCTTCAAAGCTGCTTCAGGGCGTTCAGCTGACCGGTCAAACTGGGTAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTGACAACGGTGCTTCAAAGCTGCTTCAGGGCGTTCAGCTGACCGGTCAAACTGGGTAC 300

Qy        301 ATCACTGACGATCTGGACATCTGACATCTCACCGGTTTCTACCGGTAACGGTTTCCACCGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCACTGACGATCTGGACATCTGACATCTCACCGGTTTCTACCGGTAACGGTTTCCACCGAC 360

Qy        361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGGAGCGAACAGCACTGGGCGTTTCCCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGGAGCGAACAGCACTGGGCGTTTCCCCA 420

Qy        421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC 480

Qy        481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGGTGGGTACCACCGTCCTGATAACGGCATG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGGTGGGTACCACCGTCCTGATAACGGCATG 540

Qy        541 CTGAGCCGTGGGCGTTTCCTACCGGTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAGCCGTGGGCGTTTCCTACCGGTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG 600

Qy        601 GCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTCCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy        661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy        721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTGTTGTTCGGGCTACACCGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCGTGTTGTTCGGGCTACACCGAC 780

Qy        781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy        841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGTGACAACGTGAAAGCTCGCCGCGGCTGGAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGTGACAACGTGAAAGCTCGCCGCGGCTGGAA 900

Qy        901 TCCAACCGGTTACTCGGCAACACCTGCGAACGTGAAAGCTGTGAAAGGCTACAAAGAGAAGTTGTAACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCCAACCGGTTACTCGGCAACACCTGCGAACGTGAAAGCTGTGAAAGGCTACAAAGAGAAGTTGTAACT 960

Qy        961 TGCCTGGCTCCGGATCCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAGAAGTTGTAACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGCCTGGCTCCGGATCCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAGAAGTTGTAACT 1020

Qy       1021 CAGCCGCGGGGTTAA 1035
              |||||||||||||||
Db       1021 CAGCCGCGGGGTTAA 1035
```

Query Match   100.0%;   Score 1035;   DB 6;   Length 1035;
Best Local Similarity 100.0%;   Pred. No. 1.1e-241;
Matches 1035;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT   60

Qy     61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCACGGTAACGGTTTCCAGAACAAC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCACGGTAACGGTTTCCAGAACAAC  120

Qy    121 AACGGTCCGACCCGTAACCAGCAGCTGTTCGGTGTCGGTGCGTTCGGTGGTTACCAGGTTAAC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACGGTCCGACCCGTAACCAGCAGCTGTTCGGTGTCGGTGCGTTCGGTGGTTACCAGGTTAAC  180

Qy    181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATATAAAGGCAGC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATATAAAGGCAGC  240

Qy    241 GTTGACAACGGTGCTTCAAAGCTCAGCTGCGTTCAGCTGCGGGTTACCCG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTGACAACGGTGCTTCAAAGCTCAGCTGCGTTCAGCTGCGGGTTACCCG  300

Qy    301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCGGCATGGTTTGGGCGCGCTGACTCC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATCACTGACGATCTGGACATCTACACCCGTCTCGGGCGGCGGCATGGTTTGGGCGCGCTGACTCC  360

Qy    361 AAAGGCAACTACCGCTTCTACCGGCGTTTCCCGTAGCGAACAGACACTCGTACCCG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGGCAACTACCGCTTCTACCGGCGTTTCCCGTAGCGAACAGACACTCGTACCCG  420

Qy    421 GTATTTGCTGCGGCGGTTAACGGCTGTTACTCGTGACATCGCTACCCGTCTCGAATAC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATTTGCTGCGGCGGTTAACGGCTGTTACTCGTGACATCGCTACCCGTCTCGAATAC  480

Qy    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG  540

Qy    541 CTGAGCCGGTCCGGCGTTGGTCCTACCCGTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGAGCCGGTCCGGCGTTGGTCCTACCCGTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG  600

Qy    601 GCTCCGGCTCCGGCGTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGCTGACGTTCTG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCGGCTCCGGCGTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGCTGACGTTCTG  660

Qy    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGTGCTGCTGATCAGCTGTAC  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGTGCTGCTGATCAGCTGTAC  720

Qy    721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTGTTCTGTGGGCTACCACCGAC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTGTTCTGTGGGCTACCACCGAC  780

Qy    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840

Qy    841 GACTACCTGGTTGCTAAAGGCATCCCGCCTGGCAAAATCTCCGCTCGGCGCATGGGTGAA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTACCTGGTTGCTAAAGGCATCCCGCCTGGCAAAATCTCCGCTCGGCGCATGGGTGAA  900

Qy    901 TCCAACCCGGTTACTGACAACACCTGTGACAACGTTGAAAGCTCGGCGCTGCCCTGATCGAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCAACCCGGTTACTGACAACACCTGTGACAACGTTGAAAGCTCGGCGCTGCCCTGATCGAT  960

Qy    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTGTAACT 1020
          ||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTGTAACT 1020

Qy   1021 CAGCCGGCGGGTTAA 1035
          |||||||||||||||
Db   1021 CAGCCGGCGGGTTAA 1035
```

RESULT 8
LOCUS   AX035701   1035 bp   DNA   linear   PAT 15-NOV-2000
DEFINITION Sequence 1 from Patent WO0054789.
ACCESSION   AX035701
VERSION   AX035701.1   GI:11191297
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Beck,A., Libon,C., Bonnefoy,J.Y., Corvaia,N. and N Guyen,T.
  TITLE   Bacterial membrane fractions with adjuvant effect
  JOURNAL   Patent: WO 0054789-A 1 21-SEP-2000;
            PF MEDICAMENT (FR) ; BECK ALAIN (FR) ; LIBON CHRISTINE (FR) ;
            BONNEFOY JEAN YVES (FR) ; CORVAIA NATHALIE (FR) ; GUYEN THIEN NGOC
            N (FR)
FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS   1..1035
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC16265.1"
                     /db_xref="GI:11191298"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQUGAGAPFGGYGVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTGVTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQALDLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSSKRA
                     QSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN
```

Qy   1021 CAGCCGGCGGGTTAA 1035
           ||||||||||||||
Db   1021 CAGCCGGCGGGTTAA 1035

RESULT 9
AX191406
LOCUS   AX191406   1035 bp   DNA   linear   PAT 15-AUG-2001
DEFINITION Sequence 1 from Patent WO0149705.
ACCESSION   AX191406
VERSION   AX191406.1 GI:15209619
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Baussant,T., Jeannin,P., Delneste,Y., Lawny,F. and Bonnefoy,J.Y.
  TITLE   Method for preparing a polypeptide soluble in an aqueous solvent in
            the absence of detergent
  JOURNAL   Patent: WO 0149705-A 1 12-JUL-2001;
            PIERRE FABRE MEDICAMENT (FR)
FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS   1..1035
                     /note="unnamed protein product; P40"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC51138.1"
                     /db_xref="GI:15209620"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFONNNGPT
                     RNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSREHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLVAKGIPAGKIISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPAG"
ORIGIN

Query Match   100.0%; Score 1035; DB 6; Length 1035;
Best Local Similarity   100.0%;
Matches 1035; Conservative   0; Mismatches   0; Indels   0; Gaps   0; Pred. No. 1.1e-241;

Qy   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCTACGGTAACGGGTTCCAGAACAAC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACACCGGTTCTACGGTAACGGGTTCCAGAACAAC 120

Qy   121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCTGGTGGCGTTCAGCTGGGTGCTGGT 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACGGTCCGACCCGTAACGATCAGCTGGGTGCTGGTGGCGTTCAGCTGGGTGCTGGT 180

Qy   181 CCGTACCTCCGGTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGTACCTCCGGTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy   241 GTTGACAACGGTGCTTCAAAGCTCAGGCGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTTGACAACGGTGCTTCAAAGCTCAGGCGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy   301 ATCACTGACGATCTGACATCTACACCCGTCCTTCACCGGTTCGGCGCGCTGACTCC 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCACTGACGATCTGACATCTACACCCGTCCTTCACCGGTTCGGCGCGCTGACTCC 360

Qy   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCCGGCGTTTCCCCGGCGTTTCCCCA 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCCGGCGTTTCCCCGGCGTTTCCCCA 420

Qy   421 GTATTTGCTGCGGCGGCTAGAGTGGCTGTTACTCGTGACATCGTACCCGTCCTGGAATAC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTATTTGCTGCGGCGGCTAGAGTGGCTGTTACTCGTGACATCGTACCCGTCCTGGAATAC 480

Qy   481 CAGTGGGTTAACAACATCGGCGACGCGGCACTGTGGTCAGTGGT 540
           ||||||||||||||||||||||||||||||||||||||||||||
Db   481 CAGTGGGTTAACAACATCGGCGACGCGGCACTGTGGTCAGTGCATG 540

Qy   541 CTGAGCCTGGGCGTTTCCTACCGGTTCGGTGCAGGAGAGATGTCAGTGTGTCTCCG 600
           |||||||||||||||||||||||||||||||||||||||||||||
Db   541 CTGAGCCTGGGCGTTTCCTACCGGTTCGGTGCAGGAGAGATGTCAGTGTGTCTCCG 600

Qy   601 GCTCCCGCTCCGGCGTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCTCCCGCTCCGGCGTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy   661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTGTTGTCTGGGCTACACCGAC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTGTTGTCTGGGCTACACCGAC 780

Qy   781 CGCATCGGTTCCGAAGCTTCAACCAGCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CGCATCGGTTCCGAAGCTTCAACCAGCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTTCGCTCGCGGCATGGGTGAA 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTTCGCTCGCGGCATGGGTGAA 900

Qy   901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCGCCCTGCCCTGATCGAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGCGCCCTGCCCTGATCGAT 960

Qy   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGGCGGGTTAA 1035
           ||||||||||||||
Db   1021 CAGCCGGCGGGTTAA 1035

RESULT 10
AX101162
LOCUS   AX101162   1032 bp   DNA   linear   PAT 10-APR-2001.
DEFINITION Sequence 1 from Patent WO0121203.
ACCESSION   AX101162
VERSION   AX101162.1 GI:13619998
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM Klebsiella pneumoniae
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   corva A,N. and Goestch,L.
  TITLE   Use of an outer membrane protein a of an enterobacterium associated
            with a rsv immunogenic peptide for preparing vaccines for
            intranasal administration
  JOURNAL   Patent: WO 0121203-A 1 29-MAR-2001;
            PIERRE FABRE MEDICAMENT (FR)
FEATURES   Location/Qualifiers
     source   1..1032
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS   1..>1032
                     /note="unnamed protein product; P40"
                     /codon_start=1
                     /transl_table=11

```
      /protein_id="CAC36502.1"
      /db_xref="GI:13619999"
      /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
      RNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
      DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
      QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
      VLFNFNKATLKPEGGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQLSEKRA
      QSVVDYLVVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
      YKBEVVTQPAG"
ORIGIN

Query Match       99.7%;   Score 1032;  DB 6;  Length 1032;
Best Local Similarity   100.0%;   Pred. No. 5.7e-241;
Matches 1032;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy  121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC 180

Qy  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG 300

Qy  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGCTGACTCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGCTGACTCC 360

Qy  361 AAAGGCAACTACGGTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGGCAACTACGGTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA 420

Qy  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy  481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540

Qy  541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCGGCTCCAGTGGTTGCTCCG 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCGGCTCCAGTGGTTGCTCCG 600

Qy  601 GCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG 660

Qy  661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC 720

Qy  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTTGTTCTGGGCTACACCGAC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTTGTTCTGGGCTACACCGAC 780

Qy  781 CGCATCGGTTCCGAAGCTTACAACCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CGCATCGGTTCCGAAGCTTACAACCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT 840

Qy  841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900

Qy  901 TCCAACCCGGTTACTGGCAACAACTGTGACAACGTGAAAGCTGAAAGCTGAAAGCTGAAAGCTGAT 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCAACCCGGTTACTGGCAACAACTGTGACAACGTGAAAGCTGAAAGCTGAAAGCTGAT 960

Qy  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTAAAGGCTACAAAGAAGTTGTAACT 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy 1021 CAGCCGGCGGGT 1032
        ||||||||||||
Db 1021 CAGCCGGCGGGT 1032
```

```
RESULT 11
KPOUTMEMA
LOCUS         KPOUTMEMA                1035 bp   DNA       linear   BCT  08-JUN-2004
DEFINITION    Klebsiella pneumoniae outer membrane protein A gene.
ACCESSION     AJ000998
VERSION       AJ000998.1  GI:3201544
KEYWORDS      outer membrane protein A.
SOURCE        Klebsiella pneumoniae
  ORGANISM    Klebsiella pneumoniae
              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
              Enterobacteriaceae; Klebsiella.
REFERENCE     1
  AUTHORS     Nguyen,T.N., Samuelson,P., Sterky,F., Merle-Poitte,C., Robert,A.,
              Baussant,T., Haeuw,J.F., Uhlen,M., Binz,H. and Stahl,S.
  TITLE       Direct chromosomal sequencing, gene isolation and heterologous
              expression of the outer membrane protein A of Klebsiella pneumoniae
  JOURNAL     Unpublished
REFERENCE     2  (bases 1 to 1035)
  AUTHORS     Haeuw,J.F.
  TITLE       Direct Submission
  JOURNAL     Submitted (04-AUG-1997) Haeuw J.F., Biochemistry Department, Centre
              d'Immunologie Pierre Fabre, 5 Avenue Napoleon III. Saint Julien en
              Genevois, 74160, FRANCE
FEATURES             Location/Qualifiers
     source          1..1035
                     /organism="Klebsiella pneumoniae"
                     /mol_type="genomic DNA"
                     /strain="RV 308"
                     /db_xref="taxon:573"
                     /plasmid="pVALOmpAKpn"
     CDS             1..1035
                     /codon_start=1
                     /transl_table=11
                     /product="outer membrane protein A"
                     /protein_id="CAA04450.1"
                     /db_xref="GI:3201545"
                     /db_xref="GOA:P24017"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQLSEKRA
                     QSVVDYLVVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKBEVVTQPQA"
ORIGIN

Query Match       99.5%;   Score 1030.2;  DB 1;  Length 1035;
Best Local Similarity   99.7%;   Pred. No. 1.6e-240;
Matches 1032;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
```

```
ORGANISM     unidentified
             unclassified.
REFERENCE    1  (bases 1 to 1035)
AUTHORS      Binz,H., Haeuw,J. and Svenson,S.
TITLE        IMMUNOGENIC COMPLEX, USE, METHOD OF PREPARATION THEREOF AND VACCINE
             CONTAINING SAME
JOURNAL      Patent: WO 9741888-A 3 13-NOV-1997;
             PF MEDICAMENT (FR)
COMMENT      Other publication FR 2748476 19971114.
FEATURES             Location/Qualifiers
source       1..1035
             /organism="unidentified"
             /mol_type="unassigned DNA"
             /db_xref="taxon:32644"
CDS          1..1035
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="CAB42227.1"
             /db_xref="GI:4756439"
             /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
             RNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
             DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVPAGGVEWAVTRDIATRLEY
             QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPEVATKHFTLKSD
             VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
             QSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
             YKEVVTQPQG"
ORIGIN
Query Match           99.5%;    Score 1030.2;   DB 6;    Length 1035;
Best Local Similarity 99.7%;    Pred. No. 1.6e-240;
Matches 1032;   Conservative  0;   Mismatches  3;   Indels  0;   Gaps   0;
```

```
Qy    1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT    60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT    60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC   120

Qy  121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC   180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC   180

Qy  181 CCGTACCTCGGTTCGAAATGGGTTATGACTGGCTGGCGGCCGTATGGCATATAAAGGCAGC   240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGTACCTCGGTTCGAAATGGGTTATGACTGGCTGGCGGCCGTATGGCATATAAAGGCAGC   240

Qy  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG   300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG   300

Qy  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGCTGACTCC   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGCTGACTCC   360

Qy  361 AAAGGCAACTACGGTTCTACCGGCGTGTAACACGGCAACAACACTGGCGTTCCCCA       420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGGCAACTACGGTTCTACCGGCGTGTAACACGGCAACAACACTGGCGTTCCCCA       420

Qy  421 GTATTTGCTGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCCTGGAATAC   480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTATTTGCTGCGGCGTAGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCCTGGAATAC   480

Qy  481 CAGTGGGTTAACAACATCGGCGACGGCACTGTGGGTACCCGTCCGATAACGGCATG       540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CAGTGGGTTAACAACATCGGCGACGGCACTGTGGGTACCCGTCCGATAACGGCATG       540

Qy  541 CTGAGCCTGGGCGTTTCCTACCGCCTTCGTCGGTTGTTGCTCCG                 600
        ||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTGAGCCTGGGCGTTTCCTACCGCCTTCGTCGGTTGTTGCTCCG                 600

Qy  601 GCTCCGCCGGTCCGGCTTCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG    660
```

```
RESULT 12
LOCUS       A67574             1035 bp    DNA       linear    PAT 05-MAY-1999
DEFINITION  Sequence 3 from Patent WO9741888.
ACCESSION   A67574
VERSION     A67574.1  GI:4756438
KEYWORDS    .
SOURCE      unidentified
```

```
Query Match              99.1%;     Score 1026.2;  DB 6;   Length 1356;
Best Local Similarity    99.7%;     Pred. No. 1.4e-239;
Matches 1028;  Conservative    0;   Mismatches   3;  Indels   0;  Gaps   0;

Qy    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT   60

Qy   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTCACGGTAACGGTTTCCAGAACAAC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAACTGGGTTGGTCCCAGTATCACGACACCGGTTTCTCACGGTAACGGTTTCCAGAACAAC  120

Qy  121 AACGGTCGACCCGTAACGATCAGCATCAGCTTCGGTGGTCGGTGCTTCGGTGGTTACCAGTTAAC  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACGGTCGACCCGTAACGATCAGCATCAGCTTCGGTGGTCGGTGCTTCGGTGGTTACCAGTTAAC  180

Qy  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC  240

Qy  241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGGTTCAGCTGACCGCTAAACTGGGTTACCCG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTTGACAACGGTGCTTCAAAGCTCAGGGCGGTTCAGCTGACCGCTAAACTGGGTTACCCG  300

Qy  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGTTGCTGACTCC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCGGCATGGTTTGGCGCGTTGCTGACTCC  360

Qy  361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AAAGGCAACTACGCTTCTACCGGCGTTTCCCGTAGCGAACACGACACTGGCGTTTCCCCA  420

Qy  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTATTTGCTGGCGGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGGAATAC  480

Qy  481 CAGTGGGTTAACACATCGGCGACCGCGGCGGCACTGTGGGTTACCCGTCCTGATAACGGCATG  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CAGTGGGTTAACACATCGGCGACCGCGGCGGCACTGTGGGTTACCCGTCCTGATAACGGCATG  540

Qy  541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CTGAGCCTGGGCGTTTCCTACCGCTTCGGTCAGGAAGATGCTGCACCGGTTGTTGCTCCG  600

Qy  601 GCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCTCCGGCTCCGGAAGTGGCTACCAAGCACTTCACCCTGAAGTCTGACGTTCTG  660

Qy  661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTCTGGATCAGCTGTAC  720

Qy  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGCTGTTGTTCTGGGCTACACCGAC  780

Qy  781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCAGTCCGTTGTT  840

Qy  841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA  900

Qy  901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCGCTGCCCTGATCGAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGCGCGCTGCCCTGATCGAT 960

Qy  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy 1021 CAGCCGCCGGGTTAA 1035
          ||||| ||| ||||
Db 1021 CAGCCGCAGGGCTAA 1035
```

RESULT 13

```
AX101164

LOCUS       AX101164                1356 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 3 from Patent WO0121203.
ACCESSION   AX101164
VERSION     AX101164.1  GI:13620000
KEYWORDS    synthetic construct
SOURCE      synthetic construct
  ORGANISM  artificial sequences.
REFERENCE   1
  AUTHORS   corva A.N. and Goestch,L.
  TITLE     Use of an outer membrane protein a of an enterobacterium associated
            with a rsv immunogenic peptide for preparing vaccines for
            intranasal administration
  JOURNAL   Patent: WO 0121203-A 3 29-MAR-2001;
            PIERRE FABRE MEDICAMENT (FR)
FEATURES             Location/Qualifiers
     source          1..1356
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="SEQUENCE NUCLEIQUE RESULTANT DE LA FUSION ENTRE DE
                     L'ADN DE KLEBSIELLA PNEUMONIAE (SEQ ID No. 1) ET DE L'ADN
                     DU VIRUS RESPIRATOIRE SYNCYTIAL (SEQ ID No. 8)"
     CDS             1..>1356
                     /note="unnamed protein product; P40G2Na"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC36503.1"
                     /db_xref="GI:13620001"
                     /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
                     RNDQLGAGAPFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPIT
                     DDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
                     QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD
                     VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
                     QSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
                     YKEVVTQPQGPGDPMTVKTKNTTTQTQPSKPTTKQRQNKPPNKPNNDFHFEVPNFVP
                     CSICSNNPTCWAICKRIPNKKPGKKTTKPTKPTFKTTKKDHKPQTTKKPKEVPTTKP
                     VD"
ORIGIN
```

```
Qy   361  AAAGGCAACTACGCTTCACCGGCGTTCCCGTAGCGAACACGACACTGGCGTTCCCA  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AAAGGCAACTACGCTTCACCGGCGTTCCCGTAGCGAACACGACACTGGCGTTCCCA  420

Qy   421  GTATTGCTGGCGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GTATTGCTGGCGCGTAGAGTGGGCTGTTACTCGTGACATCGCTACCCGTCTGAATAC  480

Qy   481  CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CAGTGGGTTAACAACATCGGCGACGCGGGCACTGTGGGTACCCGTCCTGATAACGGCATG  540

Qy   541  CTGAGCCTGGGCGTTTCCTACCGCTTCGGTGCGTTGGTCTGCACCGGTTGTTGTCTCCG  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  CTGAGCCTGGGCGTTTCCTACCGCTTCGGTGCGTTGGTCTGCACCGGTTGTTGTCTCCG  600

Qy   601  GCTCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCGACGTTCTG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GCTCCGGCTCCGGCTCCGGAAGTGCTACCAAGCACTTCACCCTGAAGTCGACGTTCTG  660

Qy   661  TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTTCAGCAGTCAGTCGTAC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGCAGGCTTCAGCAGTCAGTCGTAC  720

Qy   721  ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGGTTGTTCTGGGTCTACACCGAC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGGTTGTTCTGGGTCTACACCGAC  780

Qy   781  CGCATCGGTTCCGAAGCTTACAACCAGCACGTGCTGTCTGAGAAACGTGCTCCAGTTCCGTTGTT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  CGCATCGGTTCCGAAGCTTACAACCAGCACGTGCTGTCTGAGAAACGTGCTCCAGTTCCGTTGTT  840

Qy   841  GACTACCTGGTTGCTAAGGCATCCGGCTGGCAAAATCTCCGCTCGCGGCCATGGGTGAA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GACTACCTGGTTGCTAAGGCATCCGGCTGGCAAAATCTCCGCTCGCGGCCATGGGTGAA  900

Qy   901  TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGTCGCGCTGCCCTGATCGAT  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGTCGCGCTGCCCTGATCGAT  960

Qy   961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT  1020

Qy  1021  CAGCCGGCGGGT  1032
          ||||||||||||
Db  1021  CAGCCGGCGGGT  1032
```

```
Qy  1021  CAGCCGGCGGG  1031
          ||||||| ||||
Db  1021  CAGCCGCAGGG  1031
```

RESULT 14
BD222264
LOCUS   BD222264   1032 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   Use of active P40 conjugates for nasal delivery.
ACCESSION   BD222264
VERSION   BD222264.1 GI:33032034
KEYWORDS   JP 2002509897-A/1.
SOURCE   Klebsiella pneumoniae
  ORGANISM   Klebsiella pneumoniae
                Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                Enterobacteriaceae; Klebsiella.
REFERENCE   1 (bases 1 to 1032)
  AUTHORS   Andreoni,C., Rauly,I., NGguyen,T., Haeuw,J.F. and Baussant,T.
  TITLE   Use of active P40 conjugates for nasal delivery
  JOURNAL   Patent: JP 2002509897-A 1 02-APR-2002;
                PIERRE FABRE MEDICAMENT
COMMENT   OS   Klebsiella pneumoniae
                PN   JP 2002509897-A/1
                PD   02-APR-2002
                PF   26-MAR-1999 JP 2000540854
                PR   27-MAR-1998 FR   98/03814
                PI   CHRISTINE ANDREONI,ISABELLE RAULY,THIEN N'GUYEN,JEAN FRANCOIS
                PI   HAEUW,
                PI   THIERRY BAUSSANT
                PC   A61K39/39,A61K9/00,A61K38/00,A61K39/09,A61K39/095,A61K39/12,
                PC   A61K39/145,
                PC   A61K39/155,A61K39/385,A61K47/48,A61K47/04//C07K14/26,C07K19/00, PC
                C12N15/09,
                PC   A61K37/02,C12N15/00
                CC   rP40
FEATURES   Location/Qualifiers
     source   1..1032
                     /organism="Klebsiella pneumoniae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:573"
FH   Key   Location/Qualifiers
FT   CDS   (1)..(1032)
ORIGIN

Query Match   98.9%; Score 1024; DB 6; Length 1032;
Best Local Similarity   99.5%;
Matches 1027; Conservative   0; Mismatches   5; Indels   0; Gaps   0;

```
Qy    1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGAAAGCAATTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT  60

Qy   61  AAACTGGGTTGGTCCCAGTATCACGGACACCGGTAACGGTTCTACGGTAACGGTTTCCAGAACAAC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AAACTGGGTTGGTCCCAGTATCACGACACCGGTAACGGTTCTACGGTAACGGTTTCCAGAACAAC  120

Qy  121  AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGCGTTCGGTGGTTACCAGGTTAAC  180

Qy  181  CCGTACCTCGGTTTCGAAATGGTTATGACTGGCTGGCGGCCGTATGACCATATAAAGGCAGC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CCGTACCTCGGTTTCGAAATGGTTATGACTGGCTGGCGGCCGTATGACCATATAAAGGCAGC  240

Qy  241  GTTGACAACGGTGCTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  GTTGACAACGGTGCTTCAAAGCTCAGGGCGTTCAGCTGACCGCTAAACTGGGTTACCCG  300

Qy  301  ATCACTGACGATCTGACACCCGTCTGCGGGCGCATGGTTTGGCGCGCTGACTCC  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  ATCACTGACGATCTGACACATCTCACACCCGTCTGCGGGCGCATGGTTTGGCGCGCTGACTCC  360
```

```
Qy  1021  CAGCCGGCGGGT  1032
          |||| ||||||
Db  1021  CAGCCTCAGGCT  1032
```

RESULT 15
AX016041
LOCUS   AX016041   1032 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION   Sequence 1 from Patent WO9949892.
ACCESSION   AX016041
VERSION   AX016041.1 GI:10041662
KEYWORDS   .
SOURCE   Klebsiella pneumoniae
  ORGANISM   Klebsiella pneumoniae
                Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                Enterobacteriaceae; Klebsiella.
REFERENCE   1
  AUTHORS   Haeuw,J.F., N'Guyen,T., Rauly,I., Andreoni,C. and Baussant,T.
  TITLE   Use of active p40 conjugates for nasal delivery
  JOURNAL   Patent: WO 9949892-A 1 07-OCT-1999;
                HAEUW JEAN FRANCOIS (FR); GUYEN THIEN N (FR); PF MEDICAMENT (FR);
                RAULY ISABELLE (FR); ANDREONI CHRISTINE (FR); BAUSSANT THIERRY (FR)
FEATURES   Location/Qualifiers
     source   1..1032
                     /organism="Klebsiella pneumoniae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:573"
     CDS   1..>1032
                     /note="unnamed protein product; rP40"

```
          /codon_start=1
          /transl_table=11
          /protein_id="CAC07622.1"
          /db_xref="GI:11041663"
          /translation="MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPT
          RNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGARKAQGVQLTAKLGYPIT
          DLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEY
          QWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQRDAAPVVAPAPAPEVATKHFTLKSD
          VLFNFNKATLKPEGQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRA
          QSVVDYIVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG
          YKEVVTQPQA"

ORIGIN

Query Match         98.9%;   Score 1024;  DB 6;  Length 1032;
Best Local Similarity   99.5%;   Pred. No. 5.1e-239;
Matches 1027;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAGCAATTTTCGTACTGAATGCGGCTCCGAAAGATAACACCTGGTATGCAGGTGGT 60

Qy     61 AAACTGGGGTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAACTGGGGTGGTCCCAGTATCACGACACCGGTTTCTACGGTAACGGTTTCCAGAACAAC 120

Qy    121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGGCGTTCGGTGGTTACCAGGTTAAC 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACGGTCCGACCCGTAACGATCAGCTTGGTGCTGGTGGCGTTCGGTGGTTACCAGGTTAAC 180

Qy    181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGTACCTCGGTTTCGAAATGGGTTATGACTGGCTGGGCCGTATGGCATATAAAGGCAGC 240

Qy    241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCGTGACCCGTAAACTGGGTTACCCG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTGACAACGGTGCTTTCAAAGCTCAGGGCGTTCAGCGTGACCCGTAAACTGGGTTACCCG 300

Qy    301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCCGGCCATGGTTTGGCGCGCTGACTCC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATCACTGACGATCTGGACATCTACACCCGTCTGGGCCGGCCATGGTTTGGCGCGCTGACTCC 360

Qy    361 AAAGGCAACTACGGCTTCTACCGGCCGTTTCCCGTAGCGAACACGACACTGGCGTTCCCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAAGGCAACTACGGCTTCTACCGGCCGTTTCCCGTAGCGAACACGACACTGGCGTTCCCCA 420

Qy    421 GTATTTGCTGCGCGGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATTTGCTGCGCGGAGTGGGCGTGTTACTCGTGACATCGCTACCCGTCTGGAATAC 480

Qy    481 CAGTGGGTTAACAACATCGGCACGCGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGTGGGTTAACAACATCGGCACGCGGCACTGTGGGTACCCGTCCTGATAACGGCATG 540

Qy    541 CTGAGCGCTGGGGCGTTTCCTACCGGCTTCGGTCGGTCTACCAAGCACTTCACCCGTTGCTCCCG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTGAGCGCTGGGGCGTTTCCTACCGGCTTCGGTCGGTCTACCAAGCACTTCACCCGTTGCTCCCG 600

Qy    601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCCGGAAGACACGGTTCCGTCGACGTTCTG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTCCGGCTCCGGCTCCGGAAGTGGCTACCAAGCCGGAAGACACGGTTCCGTCGACGTTCTG 660

Qy    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGGCGCTTCTGGATCAGCTGTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCAACTTCAACAAAGCTACCCTGAAACCGGAAGGTCAGGCGCTTCTGGATCAGCTGTAC 720

Qy    721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTTGTTCTGGGCTACACCGAC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACTCAGCTGAGCAACATGGATCCGAAAGACGGTTCCGTTGTTCTGGGCTACACCGAC 780

Qy    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCCGTTGTT 840
          ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    781 CGCATCGGTTCCGAAGCTTACAACCAGCAGCTGTCTGAGAAACGTGCTCCGTTCGTT 840

Qy    841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACTACCTGGTTGCTAAAGGCATCCCGGCTGGCAAAATCTCCGCTCGCGGCATGGGTGAA 900

Qy    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGAAGCTCGGCCGCTGCCCTGATCGAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCAACCCGGTTACTGGCAACACCTGTGACAACGTGAAAGCTCGGAAGCTCGGCCGCTGCCCTGATCGAT 960

Qy    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCTGGCTCCGGATCGTCGTGTAGAGATCGAAGTTAAAGGCTACAAAGAAGTTGTAACT 1020

Qy   1021 CAGCCGGCGGGT 1032
          ||||| ||
Db   1021 CAGCCTCAGGCT 1032

Search completed: January 19, 2005, 15:31:33
Job time : 4775 secs
```

This Page Blank (uspto)

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:     January 19, 2005, 16:38:49 ; Search time 95 Seconds
                                          (without alignments)
                                          2083.462 Million cell updates/sec

Title:         US-09-913-772-2
Perfect score: 1823
Sequence:      1 MKAIFVLNAAPKDNTWYAGG.........DRRVEIEVKGYKEVVTQPAG 344

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:      1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters:      1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     UniProt 02:*
               1: uniprot_sprot:*
               2: uniprot_trembl:*

               Pred. No. is the number of results predicted by chance to have a
               score greater than or equal to the score of the result being printed,
               and is derived by analysis of the total score distribution.

                                       SUMMARIES

Result                Query
No.    Score    Match  Length  DB  ID            Description
-----  -------  -----  ------  --  ------------  ------------------
 1      1813    99.5    344    1   OMPA_KLEPN    P24017 klebsiella
 2      1542    84.6    350    1   OMPA_ENTAE    P09146 enterobacte
 3     1525.5   83.7    351    1   OMPA_SHIDY    P02935 shigella dy
 4      1511    82.9    346    1   OMPA_ECOLI    P02934 escherichia
 5      1511    82.9    346    2   Q6W821        Q6w821 shigella so
 6      1511    82.9    346    2   AAP74759      Aap74759 shigella
 7      1507    82.7    346    2   Q6PNM6        Q6pnm6 shigella bo
 8      1507    82.7    346    2   AAT02227      Aat02227 shigella
 9      1505    82.6    346    2   Q16J0         Q16j0 escherichia
10      1499    82.2    350    2   Q827S0        Q827s0 salmonella
11      1495    82.0    350    1   OMPA_SALTY    P02936 salmonella
12      1487    81.6    379    2   Q7UD17        Q7ud17 shigella fl
13      1473    80.8    348    2   Q83RX2        Q83rx2 shigella fl
14      1473    80.8    372    2   Q7X0I7        Q7x0i7 shigella fl
15      1466    80.4    327    2   OMPA_SERMA    P04845 serratia ma
16     1381.5   75.8    359    1   Q6QT47        Q6qt47 salmonella
17      1364    74.8    319    2   AAS18468      Aas18468 salmonell
18      1364    74.8    319    2   Q8ZG77        Q8zg77 yersinia pe
19     1325.5   72.7    353    2   Q8D056        Q8d056 yersinia pe
20     1325.5   72.7    363    2   AAS61569      Aas61569 yersinia
21     1325.5   72.7    363    2   Q6D6D4        Q6d6d4 erwinia car
22      1306    71.6    366    2   Q7N602        Q7n602 photorhabdu
23      1289    70.7    368    2   Q9RM69        Q9rm69 erwinia car
24     1262.5   69.3    367    2   Q99114        Q99114 escherichia
25     1128.5   61.9    238    1   Q47880        Q47880 escherichia
26      1128    61.0    244    2   OMPA_ESCFE    P24747 escherichia
27      1111    60.9    243    1   Q99115        Q99115 escherichia
28     1110.5   60.9    244    2   OMPA_ESCHE    P24754 escherichia
29      1106    60.7    243    1   Q99123        Q99123 enterobacte
30     1095.5   60.1    238    2   OMPA_ESCBL    Q99124 escherichia
31      1089    59.7    241    1
```

```
32     1056.5   58.0    238    1   OMPA_CITFR    P24016 citrobacter
33      1042    57.2    249    2   Q9L6I9        Q9L6I9 salmonella
34      1012    55.5    246    2   Q9L6I8        Q9L6I8 shigella fl
35       886    48.6    243    1   OMPA_SEROD    P24755 serratia od
36       849    46.6    162    2   Q6WAF8        Q6waf8 klebsiella
37       849    46.6    162    2   AAQ96089      Aaq96089 klebsiell
38       843    46.2    162    2   Q6WAF6        Q6waf6 klebsiella
39       843    46.2    162    2   Q6WAF7        Q6waf7 klebsiella
40       843    46.2    162    2   Q6WAF9        Q6waf9 klebsiella
41       843    46.2    162    2   AAQ96088      Aaq96088 klebsiell
42       843    46.2    162    2   AAQ96090      Aaq96090 klebsiell
43       843    46.2    162    2   AAQ96091      Aaq96091 klebsiell
44       835    45.8    162    2   Q6WAG0        Q6wag0 klebsiella
45       835    45.8    162    2   Q6WAG1        Q6wag1 klebsiella
```

                                       ALIGNMENTS

```
RESULT 1
OMPA_KLEPN
ID   OMPA_KLEPN     STANDARD;      PRT;    344 AA.
AC   P24017; O69435;
DT   01-MAR-1992 (Rel. 21, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Outer membrane protein A precursor (Outer membrane protein II).
GN   Name=ompA;
OS   Klebsiella pneumoniae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Klebsiella.
OX   NCBI_TaxID=573;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RV 308;
RX   MEDLINE=98192544; PubMed=9524233;
RA   Nguyen T.N., Samuelson P., Sterky F., Merle-Poitte C., Robert A.,
RA   Baussant T., Haeuw J.F., Uhlen M., Binz H., Stahl S.;
RT   "Chromosomal sequencing using a PCR-based biotin-capture method
RT   allowed isolation of the complete gene for the outer membrane protein
RT   A of Klebsiella pneumoniae.";
RL   Gene 210:93-101(1998).
RN   [2]
RP   SEQUENCE OF 93-335 FROM N.A.
RC   STRAIN=LD119;
RX   MEDLINE=92065252; PubMed=1955870;
RA   Lawrence J.G., Ochman H., Hartl D.L.;
RT   "Molecular and evolutionary relationships among enteric bacteria.";
RL   J. Gen. Microbiol. 137:1911-1921(1991).
CC   -!- FUNCTION: Required for the action of colicins K and L and for the
CC       stabilization of mating aggregates in conjugation. Serves as a
CC       receptor for a number of T-even like phages. Also acts as a porin
CC       with low permeability that allows slow penetration of small
CC       solutes (By similarity).
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   -!- SIMILARITY: Belongs to the ompA family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ000998; CAA04450.1; -.
DR   EMBL; M63355; AAA25119.1; -.
DR   HSSP; P02934; 1QJP.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
```

```
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Conjugation; Outer membrane; Phage recognition; Porin; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1      ?
FT   CHAIN         ?    344       Outer membrane protein A.
FT   TRANSMEM     15     28       Potential.
FT   TRANSMEM     48     60       Potential.
FT   TRANSMEM     63     78       Potential.
FT   TRANSMEM     90    100       Potential.
FT   TRANSMEM    104    119       Potential.
FT   TRANSMEM    140    151       Potential.
FT   TRANSMEM    157    173       Potential.
FT   TRANSMEM    179    190       Potential.
FT   DOMAIN      199    206       4 X 2 AA tandem repeats of A-P.
FT   REPEAT      199    200       1.
FT   REPEAT      201    202       2.
FT   REPEAT      203    204       3.
FT   REPEAT      205    206       4.
FT   DOMAIN      255    299       OmpA-like.
FT   DISULFID    309    321       By similarity.
FT   CONFLICT    335    335       Y -> I (in Ref. 2).
SQ   SEQUENCE   344 AA;  37061 MW;  AC88AAE3B7871B16 CRC64;

Query Match             99.5%;  Score 1813;  DB 1;  Length 344;
Best Local Similarity   100.0%;  Pred. No. 4.2e-131;
Matches  342;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVN  60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVN  60

Qy         61  PYLGFPEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADS  120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61  PYLGFPEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADS  120

Qy        121  KGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGM  180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121  KGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGM  180

Qy        181  LSLGVSYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLY  240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181  LSLGVSYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLY  240

Qy        241  TQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGE  300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241  TQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGE  300

Qy        301  SNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP  342
               ||||||||||||||||||||||||||||||||||||||||||
Db        301  SNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP  342


RESULT 2
OMPA_ENTAE
ID   OMPA_ENTAE     STANDARD;      PRT;   350 AA.
AC   P09146;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Outer membrane protein A precursor.
GN   Name=ompA;
OS   Enterobacter aerogenes (Aerobacter aerogenes).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Enterobacter.
OX   NCBI_TaxID=548;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84108348; PubMed=6363059;
```

```
RA   Braun G., Cole S.T.;
RT   "Molecular characterization of the gene coding for major outer
RT   membrane protein OmpA from Enterobacter aerogenes.";
RL   Eur. J. Biochem. 137:495-500(1983).
CC   -!- FUNCTION: Required for the action of colicins K and L and for the
CC       stabilization of mating aggregates in conjugation. Serves as a
CC       receptor for a number of T-even like phages. Also acts as a porin
CC       with low permeability that allows slow penetration of small
CC       solutes (By similarity).
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   -!- SIMILARITY: Belongs to the ompA family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X00254; CAA25062.1; -.
DR   PIR; S07222; S07222.
DR   HSSP; P02934; 1QJP.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Conjugation; Outer membrane; Phage recognition; Porin; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     21
FT   CHAIN        22    350       Outer membrane protein A.
FT   TRANSMEM     27     40       Potential.
FT   TRANSMEM     59     71       Potential.
FT   TRANSMEM     74     89       Potential.
FT   TRANSMEM    101    111       Potential.
FT   TRANSMEM    115    130       Potential.
FT   TRANSMEM    146    157       Potential.
FT   TRANSMEM    163    179       Potential.
FT   TRANSMEM    185    196       Potential.
FT   DOMAIN      205    212       4 X 2 AA tandem repeats of A-P.
FT   REPEAT      205    206       1.
FT   REPEAT      207    208       2.
FT   REPEAT      209    210       3.
FT   REPEAT      211    212       4.
FT   DOMAIN      261    305       OmpA-like.
FT   DISULFID    315    327       By similarity.
SQ   SEQUENCE   350 AA;  37575 MW;  6276C6F2F21065DA CRC64;

Query Match             84.6%;  Score 1542;  DB 1;  Length 350;
Best Local Similarity   87.5%;  Pred. No. 3e-110;
Matches  295;  Conservative   10;  Mismatches   26;  Indels    6;  Gaps    2;

Qy          6  VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF  65
               |||||||||||||||:::||||||||||:||||||:||||||||||||||||||||||||
Db         18  VAQAAPKDNTWYAGGKLGWSQFHDTGWY-NSNLNNNGPTHESQLGAGAFGGYQVNPYLGF  76

Qy         66  EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA  125
               ||||||||||||||||||:||||||||||||||||||||||||||||||:||||---- -
Db         77  EMGYDWLGRMPYKGVKVNGAFSSQAVQLTAKLGYPITDDLDIYTRLGGMVWRRADS---- 131

Qy        126  STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV  185
               |||||||||||||||||||||||||||||||||||||||||||||||||||:|||:|||:
Db        132  SNSIAGDNHDTGVSPVFAGGVEWAMTRDIATRLEYQWVNNIGDAGTVGVRPDNGMLSVGV  191

Qy        186  SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN  245
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    192 SYRFGQEDNAPVVAPAPAPAPEVTTKFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 251
              |||||||||||||||||||||  ||||||||||||||||||||||||||||||
Qy    246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
              ||||||||||||||||||||  |:||||||||||||||||||||  ||||||
Db    252 MDPKDGSAVVLGYTDRIGSEQINQKLSEKRAQSVVDYLVAKGIPANKISARGMGESDPVT 311

Qy    306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
              ||||||||||||||||||||||||||||||:|||||
Db    312 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKDVVTQP 348


RESULT 3
OMPA_SHIDY
ID   OMPA_SHIDY      STANDARD;      PRT;      351 AA.
AC   P02935;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Outer membrane protein A precursor (Outer membrane protein II).
GN   Name=ompA;
OS   Shigella dysenteriae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=622;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82221414; PubMed=6283478;
RA   Braun G., Cole S.T.;
RT   "The nucleotide sequence coding for major outer membrane protein OmpA
RT   of Shigella dysenteriae.";
RL   Nucleic Acids Res. 10:2367-2378(1982).
CC   -!- FUNCTION: Required for the action of colicins K and L and for the
CC       stabilization of mating aggregates in conjugation. Serves as a
CC       receptor for a number of T-even like phages. Also acts as a porin
CC       with low permeability that allows slow penetration of small
CC       solutes (By similarity).
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   -!- SIMILARITY: Belongs to the ompA family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; V01344; CAA24638.1; -.
DR   PIR; A03435; MMEBAD.
DR   HSSP; P02934; 1QJP.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Conjugation; Outer membrane; Phage recognition; Porin; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     21
FT   CHAIN        22    351       Outer membrane protein A.
FT   TRANSMEM     27     40       Potential.
FT   TRANSMEM     55     67       Potential.
FT   TRANSMEM     70     85       Potential.
FT   TRANSMEM     97    107       Potential.
FT   TRANSMEM    111    126       Potential.
FT   TRANSMEM    147    158       Potential.
FT   TRANSMEM    164    180       Potential.
FT   TRANSMEM    186    197       Potential.
```

```
FT   DOMAIN      206    213       4 X 2 AA tandem repeats of A-P.
FT   REPEAT      206    207       1.
FT   REPEAT      208    209       2.
FT   REPEAT      210    211       3.
FT   REPEAT      212    213       4.
FT   DOMAIN      262    306       OmpA-like.
FT   DISULFID    316    328       By similarity.
SQ   SEQUENCE   351 AA;  37741 MW;  1499AA5F5395B35B CRC64;

Query Match             83.7%;  Score 1525.5;  DB 1;  Length 351;
Best Local Similarity   84.6%;  Pred. No. 5.6e-109;
Matches  285;  Conservative  20;  Mismatches  27;  Indels  5;  Gaps  1;

Qy      6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
             |||||||||||||||||||||||||||           :|||||||||||||||||||
Db     18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI------DNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy     66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
             |||||||||:|||:::|::||::||||||||||||||||||:|||||||||||||| :|:
Db     73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDVYTRLGGMVWRADTKAHHN 132

Qy    126 STGVSRSEHDTGVSPVPAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
             ::|||:::||||||||||||||||:|:::|||||||||||||| :|||::||||||||||
Db    133 VTGESEKNHDTGVSPVFAGGVEWAITPEIATRLEYQWTNNIGDAHTIGTRPDNGLLSLGV 192

Qy    186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
             |||||| :|||||||||||||||  ::||||||||||||||||||||||||||:|||:||
Db    193 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQQALDQLYSQLSN 252

Qy    246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
             ||||||||||||||||||| :|||:|||||||||||||||||||:||||||||||:||||
Db    253 LDPKDGSAVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 312

Qy    306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
             ||||||||:||||||||||||||||||||||:|||||
Db    313 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 349


RESULT 4
OMPA_ECOLI
ID   OMPA_ECOLI      STANDARD;      PRT;      346 AA.
AC   P0A934;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Outer membrane protein A precursor (Outer membrane protein II*).
GN   Name=ompA; Synonyms=tolG, tut, con;
GN   OrderedLocusNames=b0957, z1307, ECs1041;
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=81053729; PubMed=6253901;
RA   Beck E., Bremer E.;
RT   "Nucleotide sequence of the gene ompA coding the outer membrane
RT   protein II of Escherichia coli K-12.";
RL   Nucleic Acids Res. 8:3011-3024(1980).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=81170587; PubMed=6260961;
RA   Movva N.R., Nakamura K., Inouye M.;
RT   "Gene structure of the OmpA protein, a major surface protein of
RT   Escherichia coli required for cell-cell interaction.";
RL   J. Mol. Biol. 143:317-328(1980).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
```

RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX   MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [7]
RP   SEQUENCE OF 22-346.
RC   STRAIN=K12;
RX   MEDLINE=81054820; PubMed=7001461;
RA   Chen R., Schmidmayr W., Kramer C., Chen-Schmeisser U., Henning U.;
RT   "Primary structure of major outer membrane protein II (ompA protein)
RT   of Escherichia coli K-12.";
RL   Proc. Natl. Acad. Sci. U.S.A. 77:4592-4596(1980).
RN   [8]
RP   SEQUENCE OF 22-32.
RC   STRAIN=K12 / W3110;
RA   Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
RA   Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
RA   Hochstrasser D.F.;
RL   Submitted (SEP-1994) to Swiss-Prot.
RN   [9]
RP   SEQUENCE OF 22-34.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [10]
RP   SEQUENCE OF 22-26.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=98291876; PubMed=9629924;
RA   Molloy M.P., Herbert B.R., Walsh B.J., Tyler M.I., Traini M.,
RA   Sanchez J.-C., Hochstrasser D.F., Williams K.L., Gooley A.A.;
RT   "Extraction of membrane proteins by differential solubilization for
RT   separation using two-dimensional gel electrophoresis.";

RL   Electrophoresis 19:837-844(1998).
RN   [11]
RP   MUTANTS RESISTANT TO PHAGE ENTRY.
RX   MEDLINE=84264337; PubMed=6085577;
RA   Morona R., Klose M., Henning U.;
RT   "Escherichia coli K-12 outer membrane protein (OmpA) as a
RT   bacteriophage receptor: analysis of mutant genes expressing altered
RT   proteins.";
RL   J. Bacteriol. 159:570-578(1984).
RN   [12]
RP   MUTANTS RESISTANT TO PHAGE ENTRY.
RX   MEDLINE=86033606; PubMed=3902787;
RA   Morona R., Kramer C., Henning U.;
RT   "Bacteriophage receptor area of outer membrane protein OmpA of
RT   Escherichia coli K-12.";
RL   J. Bacteriol. 164:539-543(1985).
RN   [13]
RP   PORIN ACTIVITY.
RC   STRAIN=K12;
RX   MEDLINE=92129334; PubMed=1370823;
RA   Sugawara E., Nikaido H.;
RT   "Pore-forming activity of OmpA protein of Escherichia coli.";
RL   J. Biol. Chem. 267:2507-2511(1992).
RN   [14]
RP   SUBCELLULAR LOCATION.
RX   PubMed=7813480;
RA   Kuhn A., Kiefer D., Koehne C., Zhu H.-Y., Tschantz W.R., Dalbey R.E.;
RT   "Evidence for a loop-like insertion mechanism of pro-Omp A into the
RT   inner membrane of Escherichia coli.";
RL   Eur. J. Biochem. 226:891-897(1994).
RN   [15]
RP   TOPOLOGY.
RX   MEDLINE=9148615; PubMed=8106193;
RA   Gromiha M.M., Ponnuswamy P.K.;
RT   "Prediction of transmembrane beta-strands from hydrophobic
RT   characteristics of proteins.";
RL   Int. J. Pept. Protein Res. 42:420-431(1993).
RN   [16]
RP   TOPOLOGY.
RX   MEDLINE=99295577; PubMed=10368142;
RA   Koebnik R.;
RT   "Structural and functional roles of the surface-exposed loops of the
RT   beta-barrel membrane protein OmpA from Escherichia coli.";
RL   J. Bacteriol. 181:3688-3694(1999).
RN   [17]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 22-192.
RX   MEDLINE=99023200; PubMed=9808047;
RA   Pautsch A., Schulz G.E.;
RT   "Structure of the outer membrane protein A transmembrane domain.";
RL   Nat. Struct. Biol. 5:1013-1017(1998).
RN   [18]
RP   X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RX   MEDLINE=20229895; PubMed=10764596;
RA   Pautsch A., Schulz G.E.;
RT   "High-resolution structure of the OmpA membrane domain.";
RL   J. Mol. Biol. 298:273-282(2000).
RN   [19]
RP   STRUCTURE BY NMR OF 22-197.
RX   PubMed=11276254; DOI=10.1038/86214;
RA   Arora A., Abildgaard F., Bushweller J.H., Tamm L.K.;
RT   "Structure of outer membrane protein A transmembrane domain by NMR
RT   spectroscopy.";
RL   Nat. Struct. Biol. 8:334-338(2001).
RN   [20]
RP   MASS SPECTROMETRY.
RX   MEDLINE=20222957; PubMed=10757971;
RA   le Coutre J., Whitelegge J.P., Gross A., Turk E., Wright E.M.,
RA   Kaback H.R., Faull K.F.;
RT   "Proteomics on full-length membrane proteins using mass
RT   spectrometry.";
RL   Biochemistry 39:4237-4242(2000).
CC   -!- FUNCTION: Required for the action of colicins K and L and for the
CC   stabilization of mating aggregates in conjugation. Serves as a

CC   receptor for a number of T-even like phages. Also acts as a porin
CC   with low permeability that allows slow penetration of small
CC   solutes.
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   -!- MASS SPECTROMETRY: MW=35177; METHOD=Electrospray; RANGE=22-346;
CC   NOTE=Ref.20.
CC   -!- SIMILARITY: Belongs to the ompA family.

DR   EMBL; V00307; CAA23588.1; -.
DR   EMBL; AE000198; AAC74043.1; -.
DR   EMBL; D90733; BAA35715.1; -.
DR   EMBL; D90734; BAA35722.1; -.

Query Match   82.9%; Score 1511; DB 1; Length 346;
Best Local Similarity   84.9%; Pred. No. 7.1e-108;
Matches 286; Conservative 17; Mismatches 24; Indels 10; Gaps 2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
           ||||||||||| |||||||||||||| :||||||||||||||||:||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADSKGNYA 125
           |||||||||| |||||:||| ||||||||||||||||||||||||||||||||:|:|
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADTKSN-- 130

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDNGMLSLGV 185
            : : :|||||||||||||| |:| :|||||||||:|||||:|:|||||||||||||
Db   131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy   186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
           |||||| |||||||||||||||:|||||||||||||||||||||||||:||||||:|||
Db   188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 247

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
           :|||||:|||||||||||:|||||:|||:|||||||||::||||:|||||||||||||||
Db   248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
           ||||||||:|||||||||||||||||||||:|:||||
Db   308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 5
Q6W821
ID   Q6W821   PRELIMINARY;   PRT;   346 AA.
AC   Q6W821;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Outer membrane protein A precursor.
GN   Name=ompA;
OS   Shigella sonnei.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=624;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USMS2;
RA   Kirnpal-Kaur B., Mohd Zaki S., Asma I., Ravichandran M.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY305874; AAP74759.1; -.
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR001035; MotY.

DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01023; NAFLGMOTY.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Signal.
FT   SIGNAL   1   21   Potential.
SQ   SEQUENCE   346 AA; 37201 MW; 195147734CDF8B04 CRC64;

Query Match   82.9%; Score 1511; DB 2; Length 346;
Best Local Similarity   84.9%; Pred. No. 7.1e-108;
Matches 286; Conservative 17; Mismatches 24; Indels 10; Gaps 2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
           ||||||||||| |||||||||||||| :||||||||||||||||:||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADSKGNYA 125
           |||||||||| |||||:||| ||||||||||||||||||||||||||||||||:|:|
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADTKSN-- 130

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDNGMLSLGV 185
            : : :|||||||||||||| |:| :|||||||||:|||||:|:|||||||||||||
Db   131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy   186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
           |||||| |||||||||||||||:|||||||||||||||||||||||||:||||||:|||
Db   188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 247

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
           :|||||:|||||||||||:|||||:|||:|||||||||::||||:|||||||||||||||
Db   248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
           ||||||||:|||||||||||||||||||||:|:||||
Db   308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 6
AAP74759
ID   AAP74759   PRELIMINARY;   PRT;   346 AA.
AC   AAP74759;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Outer membrane protein A precursor.
GN   OMPA.
OS   Shigella sonnei.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=624;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USMS2;
RA   Kirnpal-Kaur B., Mohd Zaki S., Asma I., Ravichandran M.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY305874; AAP74759.1; -.
KW   Signal.
FT   SIGNAL   1   21   Potential.
SQ   SEQUENCE   346 AA; 37201 MW; 195147734CDF8B04 CRC64;

Query Match   82.9%; Score 1511; DB 2; Length 346;
Best Local Similarity   84.9%; Pred. No. 7.1e-108;
Matches 286; Conservative 17; Mismatches 24; Indels 10; Gaps 2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
           ||||||||||| |||||||||||||| :||||||||||||||||:||

Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF   72
          ||||||||:|| |||||||||||||| ||||| :|||||||||||||| ||
Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          ||||||||||:|||||:|||:|||||||||||||||||||||||||||||||||:|:
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 130

Qy   126 STGVSREHDTGVSPVFAGGVEVAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
Db   131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy   186 SYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN   245
Db   188 SYRFGQGEAAPVVAPAPAPEVQTKHFTLKSDVLFNFNKATLKPGQAALDQLYSQLSN   247

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
Db   248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
Db   308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 7
Q6PNM6
ID   Q6PNM6;   PRELIMINARY;   PRT;   346 AA.
AC   Q6PNM6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Outer membrane protein A precursor.
GN   Name=ompA;
OS   Shigella boydii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=621;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MalaysiaSB631;
RA   Nik Noorul Shakira M.S., Kirnpal-Kaur B.S., Ravichandran M., Asma I.,
RA   Mohd Zaki S.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   EMBL; AY590298; AAT02227.1; -.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR001035; MotY.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01389; NAFLGMOTY.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01023; OmpA/MotB; 1.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Signal.
FT   SIGNAL   1   21   Potential.
SQ   SEQUENCE   346 AA;   37229 MW;   E94650734CDF9239 CRC64;

Query Match   82.7%;   Score 1507; DB 2;   Length 346;
Best Local Similarity 84.6%;   Pred. No. 1.4e-107;
Matches   285;   Conservative   17;   Mismatches   25;   Indels   10;   Gaps   2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 130

Qy   126 STGVSREHDTGVSPVFAGGVEVAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          |||||||||||||||||||||| :| |||||||||| :|| |||||||||||||
Db   131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy   186 SYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN   245
Db   188 SYRFGQGEAAPVVAPAPAPEVQTKHFTLKSDVLFNFNKATLKPGQAALDQLYSQLSN   247

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLIVAKGIPAGKISARGMGESNPVT 305
Db   248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPVDKISARGMGESNPVT 307

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
Db   308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 8
AAT02227
ID   AAT02227;   PRELIMINARY;   PRT;   346 AA.
AC   AAT02227;
DT   10-MAY-2004 (TrEMBLrel. 27, Created)
DT   10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT   10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE   Outer membrane protein A precursor.
GN   OMPA.
OS   Shigella boydii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=621;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MalaysiaSB631;
RA   Nik Noorul Shakira M.S., Kirnpal-Kaur B.S., Ravichandran M., Asma I.,
RA   Mohd Zaki S.;
RT   "Shigella boydii outer membrane protein A precursor (ompA) gene.";
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY590298; AAT02227.1; -.
KW   Signal.
FT   SIGNAL   1   21   Potential.
SQ   SEQUENCE   346 AA;   37229 MW;   E94650734CDF9239 CRC64;

Query Match   82.7%;   Score 1507; DB 2;   Length 346;
Best Local Similarity 84.6%;   Pred. No. 1.4e-107;
Matches   285;   Conservative   17;   Mismatches   25;   Indels   10;   Gaps   2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 130

Qy   126 STGVSREHDTGVSPVFAGGVEVAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
Db   131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy   186 SYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN   245
Db   188 SYRFGQGEAAPVVAPAPAPEVQTKHFTLKSDVLFNFNKATLKPGQAALDQLYSQLSN   247

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLIVAKGIPAGKISARGMGESNPVT 305
Db   248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPVDKISARGMGESNPVT 307

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
Db   308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 9
Q9L6J0

```
ID   Q9L6J0      PRELIMINARY;      PRT;    346 AA.
AC   Q9L6J0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Outer membrane protein A.
GN   Name=ompA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RS218;
RA   Wang Y., Kim K.S.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   EMBL; AF234269; AAF37887.1; -.
DR   HSSP; P02934; 1QJP.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01023; NAFLGMOTY.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
SQ   SEQUENCE   346 AA;  37188 MW;  5736077E41BD84C3 CRC64;

Query Match             82.6%;  Score 1505;  DB 2;  Length 346;
Best Local Similarity   84.3%;  Pred. No. 2.1e-107;
Matches  284;  Conservative   18;  Mismatches   25;  Indels   10;  Gaps    2;

Qy          6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
              ||||||||||||  ||||||||||||:|     :|||||:||||||||||||:||
Db         18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy         66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
              |||||||||| ||||:|||| |||||||||||||||||||||:|||||||||||:|.
Db         73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDVYTRLGGMVWRADTKSN-- 130

Qy        126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
                 :  ||:||||||||||||||:|:| |:|||||||||:|||||.|.|||||||||||||
Db        131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy        186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
              ||||||.:||||||||||||||| |||||||||||.|||||||||||||||||||:||||
Db        188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLTFNKATLKPEGQQALDQLYSQLSN 247

Qy        246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
              :|||||:.||||||||||:.|||| ||:|||||||||||:|||||.|||||||||||||||
Db        248 LDPKDGSVVVLGYTDRIGSDAYNQALSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy        306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
              ||||||||:||||||||||||||||||||:|:||||||
Db        308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 10
ID   Q8Z7S0      PRELIMINARY;      PRT;    350 AA.
AC   Q8Z7S0; Q7C962;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Outer membrane protein A.
GN   Name=ompA; OrderedLocusNames=STY1091, t1850;
```

```
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22511367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   EMBL; AL627269; CAD08196.1; -.
DR   EMBL; AE016840; AAO69468.1; -.
DR   HSSP; P02934; 1QJP.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR001035; MotB.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01023; NAFLGMOTY.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Complete proteome.
SQ   SEQUENCE   350 AA;  37477 MW;  EEB90059DFCADFD3 CRC64;

Query Match             82.2%;  Score 1499;  DB 2;  Length 350;
Best Local Similarity   82.8%;  Pred. No. 6e-107;
Matches  279;  Conservative   26;  Mismatches   26;  Indels    6;  Gaps    2;

Qy          6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
              ||||||||||||  ||||||||||||:|     :|||||:||||||||||||:||
Db         18 VAQAAPKDNTWYAGAKLGWSQYHDTGFI-----HNDGPTHENQLGAGAFGGYQVNPYVGF 72

Qy         66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
              |||||||||| ||||:|||| ||||||||||||||||||||:|||||||||||:|.
Db         73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDVYTRLGGMVWRADTKSNVP 132

Qy        126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
                 :  ||:||||||||||||||:|:| |:|||||||||:|||||.|.|||||||.|.:||
Db        133 G-GASTKDHDTGVSPVFAGGIEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSVGV 191

Qy        186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
              ||||||.:|||||||||||  ||| |||||||||||.|||:|||||||||||||||:||||
Db        192 SYRFGQQEAAPVVAPAPAPEVQTKHFTLKSDVLFNFNKSTLKPEGQQALDQLYSQLSN 251

Qy        246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
              :|||||:.||||||||||:.|||| ||:|||||||||||:|||||.|||||||||||||||
```

```
KW   Porin; Repeat; Signal; Transmembrane.
FT   SIGNAL         1     21
FT   CHAIN         22    350       Outer membrane protein A.
FT   TRANSMEM      27     40       Potential.
FT   TRANSMEM      55     67       Potential.
FT   TRANSMEM      85     97       Potential.
FT   TRANSMEM      97    107       Potential.
FT   TRANSMEM     111    126       Potential.
FT   TRANSMEM     146    157       Potential.
FT   TRANSMEM     163    179       Potential.
FT   TRANSMEM     185    196       Potential.
FT   DOMAIN       205    212       4 X 2 AA tandem repeats of A-P.
FT   REPEAT       205    206       1.
FT   REPEAT       207    208       2.
FT   REPEAT       209    210       3.
FT   REPEAT       211    212       4.
FT   DOMAIN       261    305       OmpA-like.
FT   DISULFID     315    327       By similarity.
FT   CONFLICT     114    114       V -> F (in Ref. 1).
FT   CONFLICT     247    247       S -> I (in Ref. 1).
SQ   SEQUENCE   350 AA;  37515 MW;  B4AC52C8C5DF54FE CRC64;

Query Match             82.0%;  Score 1495; DB 1;  Length 350;
Best Local Similarity   82.8%;  Pred. No. 1.2e-106;
Matches  279;  Conservative  26;  Mismatches  26;  Indels   6;  Gaps   2;

Qy       6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGNFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
           |||||| :|||||||||||||||||  :||||: :||||||||:::|||||||||||||||
Db      18 VAQAAPKDNTWYAGAKLGWSQYHDTGFI-----HNDGPTHENQLGAGAFGGYQVNPYVGF 72

Qy      66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
           ||||||||||| ||||| :|||||||||||||||||||||||:| ||||||||:||: ::
Db      73 EMGYDWLGRMPYKGDNINGAYKAQGVQLTAKLGYPITDDLDVYTRLGGMVWRADTKSNVP 132

Qy     126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
            ||| : :|||||||||||||||||: :|||||||||||||| ||||||||||||::|||
Db     133 G-GPSTKDHDTGVSPVFAGGIEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSVGV 191

Qy     186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
            ||||||: |||||||||||||| ||||||||||||||||:||||||||||||:||||||
Db     192 SYRFGQEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKSTLKPEGQQALDQLYSQLSN 251

Qy     246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
            :||||||::||||||||||||||||||||||||||||||:::||||||||||||||||||
Db     252 LDPKDGSVVVLGFTDRIGSDAYNQGLSEKRAQSVVDYLISKGIPSDKISARGMGESNPVT 311

Qy     306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            |||||||| |||||||||||||||||||||:|||:||
Db     312 GNTCDNVKPRAALIDCLAPDRRVEIEVKGVKDVVTQP 348
```

```
RESULT 11
OMPA_SALTY
ID   OMPA_SALTY   STANDARD;     PRT;   350 AA.
AC   P02936;
DT   21-JUL-1986 (Rel. 01, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Outer membrane protein A precursor (Outer membrane protein 33K) (Outer
DE   membrane major heat-modifiable protein).
GN   Name=ompA; OrderedLocusNames=STM1070;
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83287368; PubMed=6349993;
RA   Freudl R., Cole S.T.;
RT   "Cloning and molecular characterization of the ompA gene from
RT   Salmonella typhimurium.";
RL   Eur. J. Biochem. 134:497-502(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: Required for the action of colicins K and L and for the
CC       stabilization of mating aggregates in conjugation. Serves as a
CC       receptor for a number of T-even like phages. Also acts as a porin
CC       with low permeability that allows slow penetration of small
CC       solutes (By similarity).
CC   -!- SUBUNIT: Monomer (Probable).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC   -!- SIMILARITY: Belongs to the ompA family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X02006; CAA26037.1; -.
DR   EMBL; AE008746; AAL20003.1; -.
DR   PIR; A03436; MMEBAT.
DR   HSSP; P02934; 1QJP.
DR   StyGene; SG10263; ompA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA_membrane; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
KW   Complete proteome; Conjugation; Outer membrane; Phage recognition;
```

```
Db     252 LDPKDGSVVVLGFTDRIGSDAYNQGLSEKRAQSVVDYLISKGIPSDKISARGMGESNPVT 311

Qy     306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            |||||||| |||||||||||||||||||||:|||:||
Db     312 GNTCDNVKPRAALIDCLAPDRRVEIEVKGVKDVVTQP 348
```

```
RESULT 12
Q8CW76
ID   Q8CW76       PRELIMINARY;    PRT;   379 AA.
AC   Q8CW76;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Outer membrane protein A.
GN   Name=ompA; OrderedLocusNames=c1093;
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
```

RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
CC -!- SIMILARITY: Belongs to the ompA family.
DR EMBL; AE016758; AAN79561.1; -.
DR HSSP; P02934; 1QJP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR006664; OmpA.
DR InterPro; IPR002368; OmpA.
DR InterPro; IPR006665; OmpA/MotB.
DR InterPro; IPR006690; OMPA_LIKE.
DR InterPro; IPR000498; OmpA_tmem.
DR Pfam; PF00691; OmpA; 1.
DR Pfam; PF01389; OmpA_membrane; 1.
DR PRINTS; PR01021; OMPADOMAIN.
DR PRINTS; PR01022; OUTRMMBRANEA.
DR ProDom; PD000930; OmpA/MotB; 1.
DR PROSITE; PS01068; OMPA; 1.
KW Complete proteome.
SQ SEQUENCE   379 AA; 41054 MW; CE396D1522B1EDCA CRC64;

Query Match   81.6%; Score 1487; DB 2; Length 379;
Best Local Similarity   83.7%; Pred. No. 5, 6e-106;
Matches 282; Conservative   17; Mismatches 32; Indels   6; Gaps   2;

```
Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
         |||||||||||  ||||||||||||||      |||||    ||||||||||||||||
Db   47 VAQAAPKDNTWYTGAKLGWSQYHDTGFI----PNNGPTHENQLGAGAFGGYQVNPYVGF 101

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
         |||||||||| ||| ||||| |||||||||||||||||||||||||||||||:| |
Db  102 EMGYDWLGRMPYKGDNINGAYKAQGVLTAKLGYPITDDLDIYTRLGGMVWRADTKANVP 161

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          ||||: |:|||||||||||||  |||||||||||||| |||||| | ||||||||||||
Db  162 G-GASFKDHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 220

Qy  186 SYRFGQEDAAPVVAPAPAPAEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
         ||||||  | |||||||||||| ||||||||||| |||||||||||| ||||||: |||
Db  221 SYRFGQGEVAPVVAPAPAPAEVQTKHFTLKSDVLFTFNKATLKPEGQAALDQLYSQLSN 280

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          |||||| :|||||||||||||||| |||:|||||||||::|||| |||||||||||||||
Db  281 LDPKDGSVVLGYTDRIGSDAYNQALSERRAQSVVDYLISKGIPADKISARGMGESNPVT 340

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIRVKGYKEVVTQP 342
         |||||||: |||||||||||||||| :|| :|:| ||
Db  341 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 377
```

RESULT 13
Q7UD17
ID Q7UD17   PRELIMINARY;   PRT;   348 AA.
AC Q7UD17;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Outer membrane protein 3a (II*;G;d).
GN Name=ompA; OrderedLocusNames=S1023;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella

RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016981; AAP16471.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR InterPro; IPR006665; OmpA/MotB.
DR InterPro; IPR006690; OMPA_LIKE.
DR InterPro; IPR000498; OmpA_tmem.
DR Pfam; PF00664; Bac_OmpA.
DR Pfam; PF01389; OmpA_membrane; 1.
DR ProDom; PD000930; OmpA/MotB; 1.
DR PROSITE; PS01068; OMPA; 1.
SQ SEQUENCE   348 AA; 37283 MW; 4262C67041D62F4 CRC64;

Query Match   80.8%; Score 1473; DB 2; Length 348;
Best Local Similarity   83.4%; Pred. No. 6e-105;
Matches 281; Conservative   18; Mismatches 30; Indels   8; Gaps   3;

```
Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
         |||||||||||  ||||||||||||||      |||||    ||||||||||||||||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI----PNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
         |||||||||| ||| ||||| |||||||||||||||||||||||||||||||:| |
Db   73 EMGYDWLGRMPYKGDNINGAYKAQGVLTAKLGYPITDDLDIYTRLGGMVWRADTKANVP 132

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          ||||: |:|||||||||||||  |||||||||||||| |||||| | |||||| |||||
Db  133 G-GASFKDHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSLGV 191

Qy  186 SYRFGQEDAAPVVAPAPAPAEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
         ||||||  |||||  ||||||| |||||||||||||||||||||||| ||||||: |||
Db  192 SYRFGQGEAAPVV--APAPAEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 249

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          |||||| :|||||||||||||||| |||:|||||||||::|||| |||||||||||||||
Db  250 LDPKDGSVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 309

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIRVKGYKEVVTQP 342
         |||||||: |||||||||||||||| :|| :|:| ||
Db  310 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 346
```

RESULT 14
Q83RX2
ID Q83RX2   PRELIMINARY;   PRT;   372 AA.
AC Q83RX2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Outer membrane protein A.
GN Name=ompA; OrderedLocusNames=SF0957;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang J., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
CC -!- SIMILARITY: Belongs to the ompA family.
DR EMBL; AE015125; AAN42586.1; -.
DR HSSP; P02934; 1QJP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.

DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR002368; OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA membrane; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OmpA; 1.
KW   Complete proteome.
SQ   SEQUENCE   372 AA; 40289 MW; E5D55235E1F0CF1B CRC64;

Query Match   80.8%; Score 1473; DB 2; Length 372;
Best Local Similarity   83.4%; Pred. No. 6.5e-105;
Matches 281; Conservative 18; Mismatches 30; Indels 8; Gaps 3;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF 65
            |||||||||||||||||||||||||||| ||||:||||||||||||:|
Db   42 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----PNNGPTHENQLGAGAFGGYQVNPYVGF 96

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
            ||||||||:|||||||||||||||||||||||||||||||||||||||||:|
Db   97 EMGYDWLGRMPYKGDNINGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKANVP 156

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
            :|||||||||||||||||||:|||||||||||||||||| ||||||||||||||
Db   157 G-GASFKDHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSLGV 215

Qy   186 SYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
            ||||:||| :||||| ||||||||||||||||||||||||||| ||||:|||||
Db   216 SYRFGQGERAPVV--APAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 273

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
            ||||||||||||||||||||:||||||||||||||||||:|:||||||||||||||||||
Db   274 LDPKDGSVVVLGYTDRIGSDAYNQQLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 333

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            ||||||||:||||||||||||||||||||| |:||||
Db   334 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 370

RESULT 15
Q7X0I7
ID   Q7X0I7   PRELIMINARY;   PRT;   327 AA.
AC   Q7X0I7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Outer membrane protein A (Fragment).
GN   Name=ompA;
OS   Shigella flexneri 2a.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=42897;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USM480;
RA   Mohd Zaki S., Kirnpal-Kaur B., Ravichandran M., Asma I.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   EMBL; AY305875; AAP74760.2; -.
DR   GO; GO:0009288; C:flagellum (sensu Bacteria); IEA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR   GO; GO:0003774; F:motor activity; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   GO; GO:0001539; P:ciliary/flagellar motility; IEA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR001035; MotY.
DR   InterPro; IPR002368; OmpA.

DR   InterPro; IPR006665; OmpA/MotB.
DR   InterPro; IPR006690; OMPA_LIKE.
DR   InterPro; IPR000498; OmpA_tmem.
DR   Pfam; PF00691; OmpA; 1.
DR   Pfam; PF01389; OmpA membrane; 1.
DR   PRINTS; PR01023; NAFLGMOTY.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   PRINTS; PR01022; OUTRMMBRANEA.
DR   ProDom; PD000930; OmpA/MotB; 1.
DR   PROSITE; PS01068; OMPA; 1.
FT   NON_TER   1
SQ   SEQUENCE   327 AA; 35254 MW; 31AEFD69EF7A6182 CRC64;

Query Match   80.4%; Score 1466; DB 2; Length 327;
Best Local Similarity   83.8%; Pred. No. 1.9e-104;
Matches 279; Conservative 18; Mismatches 28; Indels 8; Gaps 3;

Qy   10 APKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGFEMGY 69
            ||||||||||||||||||||||||| ||:||||||||||||||||||
Db   1 APKDNTWYTGAKLGWSQYHDTGFI-----PNNGPTHENQLGAGAFGGYQVNPYVGFEMGY 55

Qy   70 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 129
            |||||||:||||||||||||||||||||||||||||||||||||||||:|:|
Db   56 DWLGRMPYKGDNINGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKANVPG-GA 114

Qy   130 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVVSYRF 189
            :||||||||||||||:|||||||||||||||| ||||||||||||||:
Db   115 SFKDHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSLGVVSYRF 174

Qy   190 GQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPK 249
            ||| :||||| |||||||||||||||||||||||||||| ||||:|||||:||:
Db   175 GQGEAAPVV--APAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPK 232

Qy   250 DGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTC 309
            ||||||||||||||||:|||||||||||||||||::||||||||||||||||||||||
Db   233 DGSVVVLGYTDRIGSDAYNQQLSERRAQSVVDYLISKGIPADKISARGMGESNPVTGNTC 292

Qy   310 DNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            |||||:|||||||||||||||||||| |:||||
Db   293 DNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 325

Search completed: January 19, 2005, 18:16:30
Job time : 100 secs

OM protein - protein search, using sw model

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2005  Compugen Ltd.
```

Run on: January 19, 2005, 17:58:00 ; Search time 24 Seconds
(without alignments)
1379.107 Million cell updates/sec

Title: US-09-913-772-2
Perfect score: 1823
Sequence: 1 MKAIFVLNAAPKDNTWYAGG..........DRRVEIEVKGYKEVTQPAG 344
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1782	97.8	356	2	JC6558	outer membrane pro
2	1542	84.6	350	2	S07222	outer membrane pro
3	1525.5	83.7	351	1	MMEBAD	outer membrane pro
4	1511	82.9	346	1	MMECA	outer membrane pro
5	1511	82.9	346	2	A90759	outer membrane pro
6	1511	82.9	346	2	G85622	outer membrane pro
7	1499	82.2	350	2	AI0626	outer membrane pro
8	1490	81.7	350	1	MMEBAT	outer membrane pro
9	1381.5	75.8	359	2	S07298	outer membrane pro
10	1325.5	72.7	353	2	AD0175	probable outer mem
11	1128.5	61.9	238	2	I62385	outer membrane pro
12	1111	61.0	244	2	I62393	outer membrane pro
13	1111	60.9	243	2	I62388	outer membrane pro
14	1110.5	60.9	244	2	I62389	outer membrane pro
15	1109	60.8	243	2	I84531	outer membrane pro
16	1106	60.7	243	2	I62386	outer membrane pro
17	1089	59.7	241	2	I62387	outer membrane pro
18	1084	59.5	241	2	I62391	outer membrane pro
19	1074.5	58.9	240	2	I62394	outer membrane pro
20	1056.5	58.0	238	2	I40703	outer membrane pro
21	646.5	35.5	353	2	C64187	outer membrane pro
22	515	28.3	349	2	B84968	outer membrane pro
23	443.5	24.3	321	2	F82104	outer membrane pro
24	271	14.9	194	2	A45275	21K outer membrane
25	262	14.4	236	2	A27894	outer membrane pro
26	255	14.0	242	2	C81205	outer membrane pro
27	251	13.8	240	2	A37004	outer membrane cla
28	250	13.7	242	2	A81782	outer membrane pro
29	243	13.3	326	2	S20494	root adhesin - Pse
30	231.5	12.7	350	2	S39156	outer membrane pro
31	227.5	12.5	407	2	H87341	OmpA family protei
32	208.5	11.4	319	2	C81294	outer membrane fib
33	206	11.3	344	2	A39139	outer membrane pro
34	206	11.3	389	2	B82819	outer membrane pro
35	199	10.9	210	2	A83516	probable outer mem
36	193.5	10.6	449	2	D87682	OmpA family protei
37	179	9.8	219	2	S47773	hypothetical 22.2K
38	176.5	9.7	261	2	E83185	probable outer mem
39	176	9.7	220	2	AC0982	probable outer mem
40	174	9.5	219	2	E91183	probable outer mem
41	174	9.5	219	2	A86030	probable outer mem
42	173.5	9.5	326	2	H70782	probable ompA prot
43	171.5	9.4	420	2	H95302	probable membrane
44	170	9.3	223	2	C82230	probable lipoprote
45	169.5	9.3	237	2	A83541	hypothetical prote
```

ALIGNMENTS

RESULT 1
JC6558
outer membrane protein A precursor - Klebsiella pneumoniae
C;Species: Klebsiella pneumoniae
C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 07-May-1999
C;Accession: JC6558
R;Nguyen, T.N.; Samuelson, P.; Sterky, F.; Merle-Poitte, C.; Robert, A.; Baussant, T.; H
Gene 210, 93-101, 1998
A;Title: Chromosomal sequencing using a PCR-based biotin-capture method allowed isolatio
A;Reference number: JC6558; MUID:98192544; PMID:9524233
A;Accession: JC6558
A;Molecule type: DNA
A;Residues: 1-356 <NGU>
A;Experimental source: IP I145
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-356/Product: outer membrane protein A #status predicted <MAT>

```
Query Match             97.8%;  Score 1782;  DB 2;  Length 356;
Best Local Similarity   99.4%;  Pred. No. 4.1e-127;
Matches  336;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDLGAGAFGGYQVNPYLGF 65
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 VAQAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDLGAGAFGGYQVNPYLGF 77

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   78 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 137

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  138 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 197

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGGQALDQLYTQLSN 245
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  198 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGGQALDQLYTQLSN 257

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  258 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 317

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVTQPA 343
         |||||||||||||||||||||||||||||||||||
Db  318 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVTQPA 355
```

RESULT 2
S07222

outer membrane protein ompA precursor - Enterobacter aerogenes
C;Species: Enterobacter aerogenes
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: S07222
R;Braun, G.; Cole, S.T.
Eur. J. Biochem. 137, 495-500, 1983
A;Title: Molecular characterization of the gene coding for major outer membrane protein
A;Reference number: S07222; MUID:84108348; PMID:6363059
A;Accession: S07222
A;Molecule type: DNA
A;Residues: 1-350 <BRA>
A;Cross-references: UNIPROT:P09146; EMBL:X00254; NID:g40837; PIDN:CAA25062.1; PID:g40838
A;Note: the authors translated the codon CAG for residue 197 as Asn
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-350/Product: outer membrane protein ompA #status predicted <MAT>
F;201-212/Region: alanine/proline-rich
F;213-350/Domain: periplasmic #status predicted <PER>

Query Match   84.6%;   Score 1542;   DB 2;   Length 350;
Best Local Similarity   87.5%;   Pred. No. 5.4e-109;
Matches 295; Conservative 10; Mismatches 26; Indels 6; Gaps 2;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
            |||||||||||||||||||||||||||||| ||:::||||||||||||||||||||||||
Db   18 VAQAAPKDNTWYAGGKLGWSQFHDTGWY-NSNLNNNGPTHESQLGAGAFGGYQVNPYLGF 76

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
            ||||||||| || |||||||| ||||||||||||||||||||||||||||||||||| -----
Db   77 EMGYDWLGRMPYKGVKVNGAFSSQAVQLTAKLGYPITDDLDIYTRLGGMVWRADS------ 131

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db   132 SNSIAGDNHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSVGV 191

Qy   186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
            |||||||||||||||||||||||:||| |||||||||||||||||||||||||||||||
Db   192 SYRFGQEDNAPVVAPAPAPAPEVTTKTFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 251

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 MDPKDGSAVVLGYTDRIGSEQYNQKLSEKRAQSVVDYLVAKGIPANKISARGMGESDPVT 311

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            |||||||||||||||||||||||||||||||||:|||
Db   312 GNTCDNVKARAALIDCLAPDRRVAIEVKGYKDVVTQP 348

RESULT 3
MMEBAD
outer membrane protein A precursor - Shigella dysenteriae
C;Species: Shigella dysenteriae
C;Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A03435
R;Braun, G.; Cole, S.T.
Nucleic Acids Res. 10, 2367-2378, 1982
A;Title: The nucleotide sequence coding for major outer membrane protein OmpA of Shigell
A;Reference number: A03435; MUID:82221414; PMID:6283478
A;Accession: A03435
A;Molecule type: DNA
A;Residues: 1-351 <BRA>
A;Cross-references: UNIPROT:P02935; GB:V01344; NID:g46943; PIDN:CAA24638.1; PID:g46945
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-351/Product: outer membrane protein A #status predicted <MAT>
F;22-201/Domain: intramembrane #status predicted <INT>
F;201-213/Region: alanine/proline-rich

F;214-351/Domain: periplasmic #status predicted <PER>

Query Match   83.7%;   Score 1525.5; DB 1;   Length 351;
Best Local Similarity   84.6%;   Pred. No. 9.5e-108;
Matches 285; Conservative 20; Mismatches 27; Indels 5; Gaps 1;

Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
            |||||||||||||||||||||:||||||| :||||||---||||||||||||||||||||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----DNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
            ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||| |
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDVYTRIGGMVWRADTKAHNN 132

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
            |||||:||||||||||||||||||||::||:|||||||||||||||||||||||||||||
Db   133 VTGESEKNHDTGVSPVFAGGVEWAITPEIATRLEYQWTNNIGDAHTIGTRPDNGLLSLGV 192

Qy   186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
            ||||||||||||||||||||||||:|||||||||||||||||||||||||||||:|||||
Db   193 SYRFGQGEAAPVVAPAPAPAPEVQTKGFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 252

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
            ||||||||||||||||||||||:|||||||||||||||||:::||||||:||||||||||
Db   253 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 312

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
            |||||||:|||||||||||||||||||||||:|||||
Db   313 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 349

RESULT 4
MMECA
outer membrane protein A precursor - Escherichia coli (strain K-12)
N;Alternate names: outer membrane protein II*
C;Species: Escherichia coli
C;Date: 30-Sep-1980 #sequence revision 30-Sep-1980 #text_change 09-Jul-2004
C;Accession: A93707; A92862; A93855; S50909; D64836; A03434
R;Beck, E.; Bremer, E.
Nucleic Acids Res. 8, 3011-3024, 1980
A;Title: Nucleotide sequence of the gene ompA coding the outer membrane protein II of Esc
A;Reference number: A93707; MUID:81053729; PMID:6253901
A;Accession: A93707
A;Molecule type: DNA
A;Residues: 1-346 <BEC>
A;Cross-references: UNIPROT:P02934; GB:V00307; GB:J01654; NID:g42159; PIDN:CAA23588.1; PI
A;Experimental source: strain K12
R;Movva, N.R.; Nakamura, K.; Inouye, M.
J. Mol. Biol. 143, 317-328, 1980
A;Title: Gene structure of the OmpA protein, a major surface protein of Escherichia coli
A;Reference number: A92862; MUID:81170587; PMID:6260961
A;Accession: A92862
A;Molecule type: DNA
A;Residues: 1-346 <MOV>
A;Cross-references: GB:J01654; GB:V00307; GB:V00358; NID:g146979; PIDN:AAA24231.1; PID:g1
A;Experimental source: K12, strain K802
R;Chen, R.; Schmidmayr, W.; Kramer, C.; Chen-Schmeisser, U.; Henning, U.
Proc. Natl. Acad. Sci. U.S.A. 77, 4592-4596, 1980
A;Title: Primary structure of major outer membrane protein II* (ompA protein) of Escheric
A;Reference number: A93855; MUID:81054820; PMID:7001461
A;Accession: A93855
A;Molecule type: protein
A;Residues: 22-346 <CHE>
A;Experimental source: K12, strain P400
R;Kuhn, A.; Kiefer, D.; Koehne, C.; Zhu, H.Y.; Tschantz, W.R.; Dalbey, R.E.
Eur. J. Biochem. 226, 891-897, 1994
A;Title: Evidence for a loop-like insertion mechanism of pro-Omp A into the inner membran
A;Reference number: S50909; MUID:95112855; PMID:7813480
A;Accession: S50909
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-46 <KUH>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64836
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown; translation not shown
A;Residues: 1-346 <BLAT>
A;Cross-references: GB:AE000198; GB:U00096; NID:g1787189; PIDN:AAC74043.1; PID:g1787191;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ompA; tolG; tut; con
A;Map position: 22 min
C;Function:
A;Description: required for the action of colicins K and L and for the stabilization of
  diffusion channels that allow penetration of various solutes
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein; monomer
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-346/Product: outer membrane protein A #status predicted <MAT>
F;22-196/Domain: intramembrane #status predicted <INT>
F;196-208/Region: alanine/proline-rich
F;209-346/Domain: periplasmic #status predicted <PER>
F;257-301/Domain: ompA-like domain #status predicted <OMP>

```
Query Match            82.9%;   Score 1511;  DB 1;  Length 346;
Best Local Similarity  84.9%;   Pred. No. 1.2e-106;
Matches 286;  Conservative 17;  Mismatches 24;  Indels 10;  Gaps 2;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||| ||||||||||||||     ||||||     ||||||||||||||:||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADSKGNYA 125
          |||||||||| |||||::| |:||||||||||||||||||||||||||||||:|:
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADTKSN-- 130

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          :::||:|| ||||||||||||:|:: ||||||||||:||||||| |:|||||||||||||
Db  131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||  ||||||||||||||:|||||||||||||||||||||||||| ||||||:|||
Db  188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 247

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          :||||||:|||||||||||:|||| ||::|||||||||::|||| |||||||||||||||
Db  248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVTQP 342
          ||||||||:||||||||||||||||||:||:::||
Db  308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344
```

RESULT 5
A90759
outer membrane protein 3a ECs1041 [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A90759
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A90759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <HAY>
A;Cross-references: UNIPROT:P02934; GB:BA000007; PIDN:BAB34464.1; PID:g13360501; GSPDB:C
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1041

C;Superfamily: outer membrane protein A

```
Query Match            82.9%;   Score 1511;  DB 2;  Length 346;
Best Local Similarity  84.9%;   Pred. No. 1.2e-106;
Matches 286;  Conservative 17;  Mismatches 24;  Indels 10;  Gaps 2;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||| ||||||||||||||     ||||||     ||||||||||||||:||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADSKGNYA 125
          |||||||||| |||||::| |:||||||||||||||||||||||||||||||:|:
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADTKSN-- 130

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          :::||:|| ||||||||||||:|:: ||||||||||:||||||| |:|||||||||||||
Db  131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||  ||||||||||||||:|||||||||||||||||||||||||| ||||||:|||
Db  188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 247

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          :||||||:|||||||||||:|||| ||::|||||||||::|||| |||||||||||||||
Db  248 LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVTQP 342
          ||||||||:||||||||||||||||||:||:::||
Db  308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344
```

RESULT 6
G85622
outer membrane protein 3a [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85622
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85622
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <STO>
A;Cross-references: UNIPROT:P02934; GB:AE005174; NID:g12514142; PIDN:AAG55443.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A

```
Query Match            82.9%;   Score 1511;  DB 2;  Length 346;
Best Local Similarity  84.9%;   Pred. No. 1.2e-106;
Matches 286;  Conservative 17;  Mismatches 24;  Indels 10;  Gaps 2;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||| ||||||||||||||     ||||||     ||||||||||||||:||
Db   18 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADSKGNYA 125
          |||||||||| |||||::| |:||||||||||||||||||||||||||||||:|:
Db   73 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLIYTRLGGMVWRADTKSN-- 130

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          :::||:|| ||||||||||||:|:: ||||||||||:||||||| |:|||||||||||||
Db  131 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||  ||||||||||||||:|||||||||||||||||||||||||| ||||||:|||
Db  188 SYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSN 247

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
```

```
Db    :||||||                     ||||||||  ||||||  ||||||||||||
Db    248 LDPKDGSVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307

Qy    306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
          ||||||||||||||||||||||||||||||  :|||||
Db    308 GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344

RESULT 7
AI0626
outer membrane protein A STY1091 [imported] - Salmonella enterica subsp. enterica serova
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
A;Accession: AI0626
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 , S.; Moule, S.; O'Gaora, P.
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
Nature 413, 848-852, 2001
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0626
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08196.1; PID:g16502245; GSPDB:GN00176
C;Genetics:
A;Gene: STY1091
C;Superfamily: outer membrane protein A

Query Match    82.2%;    Score 1499; DB 2;    Length 350;
Best Local Similarity    82.8%;    Pred. No. 9.5e-106;
Matches 279;    Conservative 26;    Mismatches 26;    Indels 6;    Gaps 2;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||||||:|||||||::|||||||:|||    :||||||||||||||||||||
Db   18 VAQAAPKDNTWYAGAKLGWSQYHDTGFI----HNDGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQQVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          |||||||||||:||:|::||||||||||||||||||:|||||::|||||||||||:|||||
Db   73 EMGYDWLGRMPYKGDNINGAYKAQQVQLTAKLGYPITDDLDFYTRLGGMVWRADTKSNVP 132

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          |||||||||||||||||||||||||| :||||||||||||||||||.||||||||||||||
Db  133 G-GASTKDHDTGVSPVFAGGIEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSVGV 191

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||::||||.|||||||||||.||||||||||||||||||||||||||||||:|||||
Db  192 SYRFGQEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKSTLKPEGQQALDQDYSQLSN 251

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          ::||||||||||||||||||||||||||||:||||||||:|||||||||||||||||||||
Db  252 LDPKDGSVVVLGFTDRIGSDAYNQGLSEKRAQSVVDYLISKGIPSDKISARGMGESNPVT 311

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
          ||||||||:||||||||||||||||||||||||||||
Db  312 GNTCDNVKPRAALIDCLAPDRRVEIEVKGVKDVVTQP 348

RESULT 8
MMEBAT
outer membrane protein A precursor - Salmonella typhimurium
N;Alternate names: outer membrane major heat-modifiable protein; outer membrane protein
C;Species: Salmonella typhimurium
C;Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 09-Jul-2004
C;Accession: AO3436
R;Freudl, R.; Cole, S.T.
Eur. J. Biochem. 134, 497-502, 1983
A;Title: Cloning and molecular characterization of the ompA gene from Salmonella typhimu
A;Reference number: AO3436; MUID:83287368; PMID:6349993
A;Accession: AO3436
A;Molecule type: DNA
```

```
A;Residues: 1-350 <FRE>
A;Cross-references: UNIPROT:P02936; GB:X02006; NID:g47798; PIDN:CAA26037.1; PID:g758341
C;Genetics:
A;Gene: ompA
A;Map position: 20 min
C;Function:
A;Description: required for the action of colicins and for the stabilization of mating as
A;Note: cannot serve as the receptor for the ompA-specific phages K3 and TuII
C;Superfamily: transmembrane protein
C;Keywords: outer membrane protein A
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-350/Product: outer membrane protein A #status predicted <MAT>
F;22-200/Domain: intramembrane #status predicted <INT>
F;200-212/Region: alanine/proline-rich
F;213-350/Domain: periplasmic #status predicted <PER>

Query Match    81.7%;    Score 1490; DB 1;    Length 350;
Best Local Similarity    82.8%;    Pred. No. 4.5e-105;
Matches 279;    Conservative 24;    Mismatches 28;    Indels 6;    Gaps 2;

Qy    6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||||||:|||||||::|||||||:|||    :||||||||||||||||||||
Db   18 VAQAAPKDNTWYAGAKLGWSQYHDTGFI----HNDGPTHENQLGAGAFGGYQVNPYVGF 72

Qy   66 EMGYDWLGRMAYKGSVDNGAFKAQQVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          |||||||||||:||:|::||||||||||||||||||:|||||::|||||||||||:|||||
Db   73 EMGYDWLGRMPYKGDNINGAYKAQQVQLTAKLGYPITDDLDFYTRLGGMVWRADTKSNVP 132

Qy  126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          |||||||||||||||||||||||||| :||||||||||||||||||.||||||||||||||
Db  133 G-GPSTKDHDTGVSPVFAGGIEYAITPEIATRLEYQWTNNIGDANTIGTRPDNGLLSVGV 191

Qy  186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||::||||.|||||||||||.||||||||||||||||||||||||||||||:|||||
Db  192 SYRFGQEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKSTLKPEGQQALDQDYIQLSN 251

Qy  246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          ::||||||||||||||||||||||||||||:||||||||:|||||||||||||||||||||
Db  252 LDPKDGSVVVLGFTDRIGSDAYNQGLSEKRAQSVVDYLISKGIPSDKISARGMGESNPVT 311

Qy  306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
          ||||||||:||||||||||||||||||||||||||||
Db  312 GNTCDNVKPRAALIDCLAPDRRVEIEVKGVKDVVTQP 348

RESULT 9
S07298
outer membrane protein ompA precursor - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C;Accession: S07298
R;Braun, G.; Cole, S.T.
Mol. Gen. Genet. 195, 321-328, 1984
A;Title: DNA sequence analysis of the Serratia marcescens ompA gene: implications for the
A;Reference number: S07298; MUID:85035845; PMID:6092858
A;Accession: S07298
A;Molecule type: DNA
A;Residues: 1-359 <BRA>
A;Cross-references: UNIPROT:P04845; EMBL:X00618; NID:g47242; PIDN:CAA25254.1; PID:g758300
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-359/Product: outer membrane protein ompA #status predicted <MAT>
F;22-209/Domain: intramembrane #status predicted <INT>
F;210-219/Region: alanine/proline-rich
F;220-359/Domain: periplasmic #status predicted <PER>

Query Match    75.8%;    Score 1381.5; DB 2;    Length 359;
Best Local Similarity    78.2%;    Pred. No. 7.4e-97;
Matches 269;    Conservative 24;    Mismatches 44;    Indels 7;    Gaps 4;
```

RESULT 11

I62385
outer membrane protein ompA - Escherichia vulneris (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia vulneris
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62385
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62385
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-238 <RES>
A;Cross-references: UNIPROT:Q99114; GB:M63348; NID:g146984; PIDN:AAA24233.1; PID:g146985
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;98-109/Region: alanine/proline-rich

Query Match   61.9%;   Score 1128.5; DB 2;   Length 238;
Best Local Similarity 90.1%;   Pred. No. 5.6e-78;
Matches 219; Conservative 7; Mismatches 12; Indels 5; Gaps 1;

```
Qy  93 LTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR 152
        |||||||||||||||||||||||||||||          :|||||||||||||||| |||
Db   1 LTAKLGYPITDDLDIYTRLGGMVWRADS----SNSIAGDNHDTGVSPVFAGGVEWAMTR 55

Qy 153 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPAPEVATGH 212
        |||||||||||||||||||||||||||||:|||||||||::|||||||||||||||: :
Db  56 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSVGVSYRFGQEENAPVVAPAPAPAPEVTTKT 115

Qy 213 FTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLS 272
        |||||||||||||||||||||||||||||||||||||||||||||||||||||:||:||
Db 116 FTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEQYNQKLS 175

Qy 273 EKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEV 332
        |||||||||||||||||:||||||||||:|||||||||||||||||||||||||||:|||
Db 176 EKRAQSVVDYLVAKGIPANKISARGMGESDPVTGNTCDNVKARAALIDCLAPDRRVAIEV 235

Qy 333 KGY 335
        |||
Db 236 KGY 238
```

RESULT 12

I62393
outer membrane protein ompA - Escherichia vulneris (ATCC 33822) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia vulneris
A;Variety: ATCC 33822
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62393
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62393
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-244 <RES>
A;Cross-references: UNIPROT:Q47880; GB:M63350; NID:g147000; PIDN:AAA24241.1; PID:g147001
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;104-115/Region: alanine/proline-rich

RESULT 10

AD0175
probable outer membrane porin A protein ompA [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD0175
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0175
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <KUR>
A;Cross-references: UNIPROT:Q8ZG77; GB:AL590842; PIDN:CAC90263.1; PID:g15979482; GSPDB:G
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A

Query Match   72.7%;   Score 1325.5; DB 2;   Length 353;
Best Local Similarity 75.2%;   Pred. No. 1.2e-92;
Matches 255; Conservative 28; Mismatches 28; Indels 49; Gaps 4;

```
Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQN--NNGPTRNDQLGAGAFGGYQVNPYL 63
        :|:||||||||| ||||||||||||||||||:|:   ||||||  ||||||| | |||||
Db  18 VAQAAPKDNTWYTGAKLGWSQYHDTGFYGNGYQNGIGNGPTHKDQLGAGAFLGLGYQANQYL 77

Qy  64 GFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGN 123
        |||:|||||||| ||||:||||||||||| |||:||||:|||||||||||||||||||:|
Db  78 GFELGYDWLGRMPYKGSVNNGAFKAQGVQLAAKLSYPIADDLDIYTRLGGMVWRADSKAN 137

Qy 124 YASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSL 183
        |  ||   ..|||||||:.|.|||:|.|..:||||  ::|.|||||||||:|||| ||||
Db 138 YGRTGQRLSDHDTGVSPLAAVGVEYALTKNWATRLDTYFVSNIGDAGTVGARPDNTMLSL 197

Qy 184 GVSYRFGQED-AAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQ 242
        |||||||::  |.  |||||||||:|.|||||||||||||||||:||.|||||||||||
Db 198 GVSYRFGQDDVVAP--APAPAPAPVVETKRFTLKSDVLFNFNKSTLKAEGQQALDQLYTQ 255

Qy 243 LSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESN 302
        ||.||||||:|||||||| :||: |||:||:|||||||||||:|||||.||||||||| .
Db 256 LSSMDPKDGSVVLGYTDAVGSDQYNQKLSEQRAQSVVDYLVSKGIPSDKISARGMGEAD 315

Qy 303 PVTGNTCDNVKARA--ALIDCLAPDRRVEIEVKGYKEVVTQPAG 344
         ||||||  .|. |  | |.|||||||||||||  ||:|||||:|
Db 316 AVTGNTCCYKSGRATKAQIVCLAPDRRVEIEVVKGIKDVVTQPQG 359
```

```
Qy   6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
        :|:|||||||||:||||||||||:||    :|||:||||||:||||||||  | ||||||
Db  18 VAQAAPKDNTWYTGGKLGWSQYQDTG----SIINNDGPTHKDQLGAGAFFGYQANQYLGF 73

Qy  66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
        ||||||||||| ||.::||||||||||| |||:||:| |||:|||||||:|||||:|:.
Db  74 EMGYDWLGRMPYKGDINNGAFKAQGVQLAAKLSYPVAQDLDVYTRLGGLVWRADAKGSF- 132

Qy 126 STGVSR-SEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLG 184
        . |. | |.|||||||:.. |.|:|.|..:||| |||||||||  .||||| |||:||.|
Db 133 DGGLDRASGHDTGVSPLVALGAEYAWTKNWATRMEYQWVNNIGDRETVGARPDNGLLSVG 192

Qy 185 VSYRFGQED-AAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQL 243
        |||||||:| ||||||| |||||| .|||||||||||| ||||.|||||||||||||.||
Db 193 VSYRFGQEDAAAPIVAPTPAPAPIVDTKRFTLKSDVLFGFNKANLKPEGQQALDQLYAQL 252

Qy 244 SNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNP 303
        |.:||||||:|||| :||||  | |  || :|||| |||||:||||| ||:||| | .||
Db 253 SSIDPKDGSVVLGFADRIGQPAPNLALSQRRADSVRDYLVSKGIPADKITARGEQANP 312

Qy 304 VTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
        ||||||||||.|||||.||||||||||||  |||||||||
Db 313 VTGNTCDNVKPRAALIECLAPDRRVEIEVVKGYKEVVTQP 351
```

```
Query Match
Best Local Similarity  61.0%;   Score 1112.5;  DB 2;  Length 244;
Matches 212;  Conservative 15;  Mismatches 15;  Indels 1;  Gaps 1;

Qy   93 LTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR 152
         ||||||||||||||||||||||||:||    ||   :||||||||||||||||||||:||
Db    1 LTAKLGYPITDDLDIYTRLGGMVWRADAKAQVPGTGASFKDHDTGVSPVFAGGVEWAMTR  60

Qy  153 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFG-QEDAAPVVAPAPAPAPEVATK 211
         |||||||||||||||||  ::||| |||||:||||||| :|||:||||||||| ::|::|
Db   61 DIATRLEYQWVNNIGDAKTLGTRQDNGMLSVGVSYRFGQQEDVAPVVAPAPAPQ-VSTK  120

Qy  212 HFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQL 271
         |||||||||||||||||||||||||||:|:|||| |||||:|||||:|||||:||||| |
Db  121 HFTLKSDVLFNFNKATLKPEGQQALDQMYSQLSNLDPKDGSVVVLGFTDRIGSDAYNQGL 180

Qy  272 SEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIE 331
         |||||||||||| :||||: :|||||||||||||||||||||| :|||||||||||||||
Db  181 SEKRAQSVVDYLLISKGIPSNKISARGMGESNPVTGNTCDNVKARPALIDCLAPDRRVEIE 240

Qy  332 VKG 334
         |||
Db  241 VKG 243                                                  g146999
```

RESULT 13
I62388
outer membrane protein A - Escherichia fergusonii (ATCC 35472) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia fergusonii
A;Variety: ATCC 35472
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62388, I62392
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
   J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62388
A;Molecule type: DNA
A;Cross-references: UNIPROT:P24747; GB:M63352; NID:g146990; PIDN:AAA24236.1; PID:g146990
A;Experimental source: ATCC 35471
A;Accession: I62392
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Cross-references: GB:M63353; NID:g146998; PIDN:AAA24240.1; PID:g146998
A;Experimental source: ATCC 35472
A;Residues: 1-243 <RE2>
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;103-114/Region: alanine/proline-rich

```
Query Match
Best Local Similarity  60.9%;   Score 1111;  DB 2;  Length 243;
Matches 211;  Conservative 13;  Mismatches 18;  Indels 0;  Gaps 0;

Qy   93 LTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR 152
         ||||||||||||||||||||||||||:||    :   ::|||||||||||||||||| |
Db    1 LTAKLGYPITDDLDIYTRLGGMVWRADTKAHNNVTGESEKNHDTGVSPVFAGGVEWAITP  60

Qy  153 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPAPEVATKH 212
         :|||||||||:||||| :|:||||||||:||||||||||::|||||||||||| ::|||
Db   61 EIATRLEYQWTNNIGDANTIGTRPDNGLLSLGVSYRFGQGEAAPVVAPAPAPE-VQTKH  120

Qy  213 FTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLS 272
         |||||||||||||||||||||:||||||:|||| |||||:||||||||||||:||||| |
Db  121 FTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKDGSVVVLGYTDRIGSDAYNQGLS 180

Qy  273 EKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEV 332
```

```
Db  181 ERRAQSVVDYLISKGIPADKISARGMGESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEV 240

Qy  333 KG 334
         ||
Db  241 KG 242
```

RESULT 14
I62389
outer membrane protein ompA - Escherichia vulneris (ATCC 33821) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia vulneris
A;Variety: ATCC 33821
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I62389
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
   J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I62389
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-244 <RES>
A;Cross-references: UNIPROT:Q99115; GB:M63349; NID:g146992; PIDN:AAA24237.1; PID:g146993
A;Experimental source: ATCC 33821
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;104-115/Region: alanine/proline-rich

```
Query Match
Best Local Similarity  60.9%;   Score 1110.5;  DB 2;  Length 244;
Matches 211;  Conservative 16;  Mismatches 15;  Indels 1;  Gaps 1;

Qy   93 LTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR 152
         ||||||||||||||||||||||||||:||    :    :||||||||||||||||||:||
Db    1 LTAKLGYPITDDLDIYTRLGGMVWRADADKSQVGGTGASFKDHDTGVSPVFAGGVEWAMTR  60

Qy  153 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFG-QEDAAPVVAPAPAPAPEVATK 211
         |||||||||||||||  ::||| |||||:||||||||| :|||:||||||||| ::|::|
Db   61 DIATRLEYQWVNNIGDAKTLGTRQDNGMLSVGVSYRFGQQEDVAPVVAPAPAPQ-VSTK  120

Qy  212 HFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQL 271
         |||||||||||||||||||||||||||:|:|||| |||||:|||||:|||||:||||| |
Db  121 HFTLKSDVLFNFNKATLKPEGQQALDQMYSQLSNLDPKDGSVVVLGFTDRIGSDAYNQGL 180

Qy  272 SEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIE 331
         |||||||||||| :||||: :|||||||||||||||||||||| :|||||||||||||||
Db  181 SEKRAQSVVDYLLISKGIPSNKISARGMGESNPVTGNTCDNVKARPALIDCLAPDRRVEIE 240

Qy  332 VKG 334
         |||
Db  241 VKG 243
```

RESULT 15
I84531
outer membrane protein A - Escherichia fergusonii (ATCC 35469) (fragment)
N;Alternate names: outer membrane protein II
C;Species: Escherichia fergusonii
A;Variety: ATCC 35469
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
C;Accession: I84531
R;Lawrence, J.G.; Ochman, H.; Hartl, D.L.
   J. Gen. Microbiol. 137, 1911-1921, 1991
A;Title: Molecular and evolutionary relationships among enteric bacteria.
A;Reference number: I40701; MUID:92065252; PMID:1955870
A;Accession: I84531
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-243 <RES>
A;Cross-references: UNIPROT:P24747; GB:M63351; NID:g146982; PIDN:AAA24232.1; PID:g146983

A;Experimental source: ATCC 35469
C;Genetics:
A;Gene: ompA
C;Superfamily: outer membrane protein A
C;Keywords: membrane protein
F;103-114/Region: alanine/proline-rich

Query Match   60.8%; Score 1109; DB 2; Length 243;
Best Local Similarity   87.2%; Pred. No. 1.7e-76;
Matches 211; Conservative   13; Mismatches   18; Indels   0; Gaps   0;

Qy   93 LTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR 152
         ||||||||||||||||||||||||||| : | — |||||||||||||||||||:|
Db   1 LTAKLGYPITDDLDIYTRLGGMVWRADTKAHNNVTGESDKNHDTGVSPVFAGGVEWAITP 60

Qy   153 DIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKH 212
         :|||||||| ||||||| |:||||||||||||||||| :|||||||||||||| |||
Db   61 EIATRLEYQWTNNIGDANTIGTRPDNGLLSLGVSYRFGQGEAAPVVAPAPAPEVQTKH 120

Qy   213 FTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLS 272
         |||||||||||||||||||||:|||||||||::||:||| |||||||||::||:|||||
Db   121 FTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKDGSVVVLGYTDRIGSDAYNQGLS 180

Qy   273 EKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEV 332
         |:||||||||| ::|||||||||||||||||||||||||||| |||||||||||||||
Db   181 ERRAQSVVDYLISKGIPADKISARGMGESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEV 240

Qy   333 KG 334
         ||
Db   241 KG 242

Search completed: January 19, 2005, 18:16:59
Job time : 27 secs

This Page Blank (uspto)

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: January 19, 2005, 18:16:43 ; Search time 526 Seconds
(without alignments)
236.281 Million cell updates/sec

Title: US-09-913-772-2
Perfect score: 1823
Sequence: 1 MKAIFVLNAAPKDNTWYAGG.........DRRVEIEVKGYKEVVTQPAG 344
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1608061 seqs, 361289386 residues
Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1823	100.0	344	14	US-10-169-953-2	Sequence 2, Appli
2	1511	82.9	346	16	US-10-467-421-90	Sequence 90, Appl
3	1511	82.9	346	16	US-10-467-421-97	Sequence 97, Appl
4	754	41.4	190	16	US-10-416-708A-24	Sequence 24, Appl
5	754	41.4	192	16	US-10-416-708A-27	Sequence 27, Appl
6	748	41.0	194	16	US-10-416-708A-10	Sequence 10, Appl
7	748	41.0	194	16	US-10-416-708A-64	Sequence 64, Appl
8	689	37.8	344	14	US-10-336-840-9	Sequence 9, Appli
9	688	37.7	344	14	US-10-336-840-6	Sequence 6, Appli
10	687	37.7	137	15	US-10-432-056-2	Sequence 2, Appli
11	687	37.7	153	9	US-09-905-176-22	Sequence 22, Appl
12	685	37.6	344	14	US-10-336-840-3	Sequence 3, Appli
13	684	37.5	344	14	US-10-336-840-1	Sequence 1, Appli
14	684	37.5	344	14	US-10-336-840-2	Sequence 2, Appli
15	684	37.5	344	14	US-10-336-840-4	Sequence 4, Appli
16	684	37.5	344	14	US-10-336-840-7	Sequence 7, Appli
17	683	37.5	344	14	US-10-336-840-10	Sequence 10, Appl
18	680.5	37.3	341	14	US-10-336-840-5	Sequence 5, Appli
19	680.5	37.3	341	14	US-10-336-840-8	Sequence 8, Appli
20	680.5	37.3	341	14	US-10-336-840-11	Sequence 11, Appl
21	678	37.2	344	14	US-10-336-840-12	Sequence 12, Appl
22	671	36.8	364	11	US-09-809-665A-151	Sequence 151, App
23	671	36.8	364	17	US-10-854-299-151	Sequence 151, App
24	667	36.6	364	17	US-10-770-824-8	Sequence 8, Appli
25	649.5	35.6	353	16	US-10-467-421-21	Sequence 21, Appl
26	646.5	35.5	353	14	US-10-203-942-9	Sequence 9, Appli
27	646.5	35.5	369	11	US-09-809-665A-153	Sequence 153, App
28	646.5	35.5	369	17	US-10-854-299-153	Sequence 153, App
29	643.5	35.3	352	14	US-10-336-840-37	Sequence 37, Appl
30	639.5	35.1	369	17	US-10-770-824-10	Sequence 10, Appl
31	506	27.8	356	14	US-10-336-840-38	Sequence 38, Appl
32	262	14.4	236	16	US-10-467-421-92	Sequence 92, Appl
33	262	14.4	236	16	US-10-467-421-96	Sequence 96, Appl
34	262	14.4	236	16	US-10-467-534-99	Sequence 99, Appl
35	255	14.0	242	16	US-10-467-421-8	Sequence 8, Appli
36	251	13.8	240	16	US-10-467-421-91	Sequence 91, Appl
37	251	13.8	240	16	US-10-467-421-95	Sequence 95, Appl
38	242	13.3	272	15	US-10-406-686A-13	Sequence 13, Appl
39	176.5	9.7	261	15	US-10-389-647-584	Sequence 584, App
40	176	9.7	390	14	US-10-283-024-3	Sequence 3, Appli
41	174.5	9.6	229	15	US-10-282-122A-69451	Sequence 69451, A
42	169.5	9.3	375	14	US-10-323-069A-123	Sequence 123, App
43	169	9.3	179	14	US-10-080-113-2	Sequence 2, Appli
44	168	9.2	395	14	US-10-323-069A-128	Sequence 128, App
45	167.5	9.2	375	14	US-10-323-069A-120	Sequence 120, App

ALIGNMENTS

RESULT 1
US-10-169-953-2

; Sequence 2, Application US/10169953
; Publication No. US2003004915A1
; GENERAL INFORMATION:
; APPLICANT: Thierry BAUSSANT
; APPLICANT: Pascale JEANNIN
; APPLICANT: Yves DELNESTE
; APPLICANT: FranOois LAWNY
; APPLICANT: Jean-Yves BONNEFOY
; TITLE OF INVENTION: METHOD FOR PREPARING A POLYPEPTIDE SOLUBLE IN AN AQUEOUS SOLVENT
; TITLE OF INVENTION: IN THE ABSENCE OF DETERGENT
; FILE REFERENCE: D18390
; CURRENT APPLICATION NUMBER: US/10/169,953
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: FR 00 00070
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; OTHER INFORMATION: P40
US-10-169-953-2

Query Match 100.0%; Score 1823; DB 14; Length 344;
Best Local Similarity 100.0%; Pred. No. 3e-160;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAPGGYQVN 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAPGGYQVN 60

Qy 61 PYILGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADS 120

```
Db    61  PYLGFEMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADS 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  KGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGM 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  KGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGM 180
Qy   181  LSLGVSYRFGQEDAAPVVAPAPAPAPEVATKGHFTLKSDVLFNFNKATLKPEGQQALDQLY 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LSLGVSYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLY 240
Qy   241  TQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGE 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGE 300
Qy   301  SNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 344
          ||||||||||||||||||||||||||||||||||||||||||||
Db   301  SNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 344
```

RESULT 2
US-10-467-421-90
; Sequence 90, Application US/10467421
; Publication No. US20040116665A1
; GENERAL INFORMATION:
;   APPLICANT: Berthet, Francois-Xavier Jacques
;   APPLICANT: Denoel, Philippe
;   APPLICANT: Neyt, Cecile Anne
;   APPLICANT: Poolman, Jan
;   APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45259
; CURRENT APPLICATION NUMBER: US/10/467,421
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01361
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103171.5
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
;   LENGTH: 346
;   TYPE: PRT
;   ORGANISM: E. coli
US-10-467-421-90

```
Query Match             82.9%;  Score 1511; DB 16;  Length 346;
Best Local Similarity   84.9%;  Pred. No. 2.6e-131;
Matches  286;  Conservative  17;  Mismatches  24;  Indels  10;  Gaps   2;

Qy     6  VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||||||||:||::||||||||  :||||||::||||||||||||||:|||
Db    18  VAQAAPKDNTWYTGAKLGWSQYHDTGFI----NNNGPTHENQLGAGAFGGYQVNPYVGF 72
Qy    66  EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          ||||||||||||||||:||||||||||||||||||||||||||||||||||||:|:|
Db    73  EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 130
Qy   126  STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          ||||||||||||||||||||||||:|:: :||||||||||||:|||||||||||||||||
Db   131  ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187
Qy   186  SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||:|||::|:|:||:||||||||||||||||||||||::|||:|||||||:||||
Db   188  SYRFGQEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQQAALDQLYSQLSN 247
Qy   246  MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          ||||||| |||::|||||||||||:|||||||||||||||: :|||||||:||||||||||
Db   248  LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307
Qy   306  GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
          ||||||:|:||||||||||||||||||||||:|:||||
Db   308  GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344
```

RESULT 3
US-10-467-421-97
; Sequence 97, Application US/10467421
; Publication No. US20040116665A1
; GENERAL INFORMATION:
;   APPLICANT: Berthet, Francois-Xavier Jacques
;   APPLICANT: Denoel, Philippe
;   APPLICANT: Neyt, Cecile Anne
;   APPLICANT: Poolman, Jan
;   APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45259
; CURRENT APPLICATION NUMBER: US/10/467,421
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: PCT/EP02/01361
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103171.5
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97
;   LENGTH: 346
;   TYPE: PRT
;   ORGANISM: E.coli
US-10-467-421-97

```
Query Match             82.9%;  Score 1511; DB 16;  Length 346;
Best Local Similarity   84.9%;  Pred. No. 2.6e-131;
Matches  286;  Conservative  17;  Mismatches  24;  Indels  10;  Gaps   2;

Qy     6  VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGF 65
          |||||||||||||||||:||::||||||||  :||||||::||||||||||||||:|||
Db    18  VAQAAPKDNTWYTGAKLGWSQYHDTGFI----NNNGPTHENQLGAGAFGGYQVNPYVGF 72
Qy    66  EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          ||||||||||||||||:||||||||||||||||||||||||||||||||||||:|:|
Db    73  EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 130
Qy   126  STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          ||||||||||||||||||||||||:|:: :||||||||||||:|||||||||||||||||
Db   131  ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGV 187
Qy   186  SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          ||||||:|||::|:|:||:||||||||||||||||||||||::|||:|||||||:||||
Db   188  SYRFGQEAAPVVAPAPAPAPEVQTKHFTLKSDVLFNFNKATLKPEGQQAALDQLYSQLSN 247
Qy   246  MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          ||||||| |||::|||||||||||:|||||||||||||||: :|||||||:||||||||||
Db   248  LDPKDGSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMGESNPVT 307
Qy   306  GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
          ||||||:|:||||||||||||||||||||||:|:||||
Db   308  GNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQP 344
```

RESULT 4
US-10-416-708A-24
; Sequence 24, Application US/10416708A
; Publication No. US20040161753A1
; GENERAL INFORMATION:
;   APPLICANT: Wise, John G.
;   APPLICANT: Fromknecht, Katja
; TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
; TITLE OF INVENTION: SPECIFICITIES
; FILE REFERENCE: 37779-0004
; CURRENT APPLICATION NUMBER: US/10/416,708A
; CURRENT FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
;   LENGTH: 190

Left column

```
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Synthetic Construct
US-10-416-708A-24

Query Match     41.4%;     Score 754; DB 16;   Length 190;
Best Local Similarity  79.9%;   Pred. No. 1.7e-61;
Matches 139;   Conservative  9;   Mismatches  16;   Indels  10;   Gaps   2;

Qy     6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF  65
           |||||| |||| ||||||||||||  |||| ||  :|||||:||:|:|
Db    20 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF  74

Qy    66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
           |||||||||| ||||:||||| ||||||||||||||||||||||||||||:|
Db    75 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 132

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDNG 179
           |||||||| ||||||||:|:| :|||||||||| ||:||||||
Db   133 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNG 183

RESULT 5
US-10-416-708A-27
; Sequence 27, Application US/10416708A
; Publication No. US20040161753A1
; GENERAL INFORMATION:
;    APPLICANT: Wise, John G.
;    APPLICANT: Fromknecht, Katja
;    TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
;    TITLE OF INVENTION: SPECIFICITIES
;    FILE REFERENCE: 37779-0004
;    CURRENT APPLICATION NUMBER: US/10/416,708A
;    CURRENT FILING DATE: 2004-01-28
;    NUMBER OF SEQ ID NOS: 89
;    SOFTWARE: PatentIn version 3.2
;    SEQ ID NO 27
;    LENGTH: 192
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Synthetic Construct
US-10-416-708A-27

Query Match     41.4%;     Score 754; DB 16;   Length 192;
Best Local Similarity  79.9%;   Pred. No. 1.7e-61;
Matches 139;   Conservative  9;   Mismatches  16;   Indels  10;   Gaps   2;

Qy     6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF  65
           |||||| |||| ||||||||||||  |||| ||  :|||||:||:|:|
Db    20 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF  74

Qy    66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
           |||||||||| ||||:||||| ||||||||||||||||||||||||||||:|
Db    75 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 132

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDNG 179
           |||||||| ||||||||:|:| :|||||||||| ||:||||||
Db   133 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDNG 183

RESULT 6
US-10-416-708A-10
; Sequence 10, Application US/10416708A
; Publication No. US20040161753A1
; GENERAL INFORMATION:
;    APPLICANT: Wise, John G.
;    APPLICANT: Fromknecht, Katja
;    TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
;    TITLE OF INVENTION: SPECIFICITIES
;    FILE REFERENCE: 37779-0004
;    CURRENT APPLICATION NUMBER: US/10/416,708A
```

Right column

```
;    CURRENT FILING DATE: 2004-01-28
;    NUMBER OF SEQ ID NOS: 89
;    SOFTWARE: PatentIn version 3.2
;    SEQ ID NO 10
;    LENGTH: 194
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Synthetic Construct
US-10-416-708A-10

Query Match     41.0%;     Score 748; DB 16;   Length 194;
Best Local Similarity  79.8%;   Pred. No. 6.4e-61;
Matches 138;   Conservative  9;   Mismatches  16;   Indels  10;   Gaps   2;

Qy     6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF  65
           |||||| |||| ||||||||||||  |||| ||  :|||||:||:|:|
Db    20 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF  74

Qy    66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
           |||||||||| ||||:||||| ||||||||||||||||||||||||||||:|
Db    75 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 132

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDN 178
           |||||||| ||||||||:|:| :|||||||||| ||:||||||
Db   133 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDN 182

RESULT 7
US-10-416-708A-64
; Sequence 64, Application US/10416708A
; Publication No. US20040161753A1
; GENERAL INFORMATION:
;    APPLICANT: Wise, John G.
;    APPLICANT: Fromknecht, Katja
;    TITLE OF INVENTION: CREATION AND IDENTIFICATION OF PROTEINS HAVING NEW DNA BINDING
;    TITLE OF INVENTION: SPECIFICITIES
;    FILE REFERENCE: 37779-0004
;    CURRENT APPLICATION NUMBER: US/10/416,708A
;    CURRENT FILING DATE: 2004-01-28
;    NUMBER OF SEQ ID NOS: 89
;    SOFTWARE: PatentIn version 3.2
;    SEQ ID NO 64
;    LENGTH: 194
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Synthetic Construct
US-10-416-708A-64

Query Match     41.0%;     Score 748; DB 16;   Length 194;
Best Local Similarity  79.8%;   Pred. No. 6.4e-61;
Matches 138;   Conservative  9;   Mismatches  16;   Indels  10;   Gaps   2;

Qy     6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGF  65
           |||||| |||| ||||||||||||  |||| ||  :|||||:||:|:|
Db    20 VAQAAPKDNTWYTGAKLGWSQYHDTGFI-----NNNGPTHENQLGAGAFGGYQVNPYVGF  74

Qy    66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
           |||||||||| ||||:||||| ||||||||||||||||||||||||||||:|
Db    75 EMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTKSN-- 132

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWNNIGDAGTVGTRPDN 178
           |||||||| ||||||||:|:| :|||||||||| ||:||||||
Db   133 ---VYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGDAHTIGTRPDN 182

RESULT 8
US-10-336-840-9
; Sequence 9, Application US/10336840
; Publication No. US20030219454A1
; GENERAL INFORMATION:
;    APPLICANT: TERRY, TAMSIN DEBORAH
```

```
;   APPLICANT:  TSENG, HSING-JU
;   APPLICANT:  HOBB, RHONDA IVY
;   APPLICANT:  JENNINGS, MICHAEL PAUL
;   APPLICANT:  DOWNES, JOHN
;   TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
;   FILE REFERENCE: 37955-0007
;   CURRENT APPLICATION NUMBER: US/10/336,840
;   CURRENT FILING DATE: 2003-01-06
;   PRIOR APPLICATION NUMBER: PCT/AU01/00822
;   PRIOR FILING DATE: 2001-07-06
;   NUMBER OF SEQ ID NOS: 39
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 9
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Haemophilus paragallinarum
US-10-336-840-9

Query Match       37.8%;   Score 689;  DB 14;  Length 344;
Best Local Similarity   44.9%;  Pred. No. 4.2e-55;
Matches   153;  Conservative  46;  Mismatches 124;  Indels 18;  Gaps  7;

Qy    1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVN  60
            |:|||  |:|:||: ||  |   :|| :| || ||::||       ::|||||
Db   13  LTAASVAQAAPQANTFYAGAKAGWASFHDGLNQFENSQNADGTLRN-SVTYGVFGGYQIT  71

Qy   61  PYLGFEMGYDWLGRMAYK-GSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRAD 119
            :|||| || ||  |:   ||| |  :||: ||  |   ||:| |  |: |:|
Db   72  DNFAVELGYDDFGRAKRRQGGETVIKYTNHGAHLSLKASYPVLEGLDVYARVGAALIRSD 131

Qy  120  SKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR--DIATRLEYQWVNNIG----DAGTVG 173
                      :  |||||  :||||||||| ::
Db  132  YKPTKRAAPNQTHEHSLKVSVFAGGLEYNLPSLPELALRVEYQWVNKVGRVEKDGSRVD  191

Qy  174  TRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFKATLKPEGQ    233
            :::  |:||||||| :|||||
Db  192  YTPSIGSVTAGLSYRFGQ--SAPVV------EPKVVAKTFALNSDVTFAFGKANLRPEAQ 243

Qy  234  QALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKI 293
            |:  | |  :::|| |||||||||||| :|:||||:||:|::::
Db  244  NVLDGIYGEIAQL--KSVQVDVAGYTDRIGSEAANLKLSQRRADTVANYLVSKGVAQEVI 301

Qy  294  SARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG 334
            |||  ||||| ||  |||:||| |||||||||||
Db  302  SSTGYGEANPVTGAKCDTVKGRKALIACLADDRRVEISVKG 342
```

RESULT 9
US-10-336-840-6

```
;   Sequence 6, Application US/10336840
;   Publication No. US20030219454A1
;   GENERAL INFORMATION:
;   APPLICANT:  TERRY, TAMSIN DEBORAH
;   APPLICANT:  TSENG, HSING-JU
;   APPLICANT:  HOBB, RHONDA IVY
;   APPLICANT:  JENNINGS, MICHAEL PAUL
;   APPLICANT:  DOWNES, JOHN
;   TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
;   FILE REFERENCE: 37955-0007
;   CURRENT APPLICATION NUMBER: US/10/336,840
;   CURRENT FILING DATE: 2003-01-06
;   PRIOR APPLICATION NUMBER: PCT/AU01/00822
;   PRIOR FILING DATE: 2001-07-06
;   NUMBER OF SEQ ID NOS: 39
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 6
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Haemophilus paragallinarum
US-10-336-840-6

Query Match       37.7%;   Score 688;  DB 14;  Length 344;
Best Local Similarity   44.9%;  Pred. No. 5.3e-55;
```

```
Matches   153;  Conservative  46;  Mismatches 124;  Indels 18;  Gaps  7;

Qy    1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVN  60
            |:|||  |:|:||: ||  |   :|| :| || ||::||       ::|||||
Db   13  LTAASVAQAAPQANTFYAGAKAGWASFHDGLNQFENSQNADGTLRN-SVTYGVFGGYQIT  71

Qy   61  PYLGFEMGYDWLGRMAYK-GSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRAD 119
            :|||| || ||  |:   ||| |  :||: ||  |   ||:| |  |: |:|
Db   72  DNFAVELGYDDFGRAKRRQGGETVIKYTNHGAHLSLKASYPVLEGLDVYARVGAALIRSD 131

Qy  120  SKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR--DIATRLEYQWVNNIG----DAGTVG 173
                      :  |||||  :||||||||| ::
Db  132  YKPTKRAAPNQTHEHSLKVSVFAGGLEYNLPSLPELALRVEYQWVNKVGRVEKDGSRVD  191

Qy  174  TRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFKATLKPEGQ    233
            :::  |:||||||| :|||||
Db  192  YTPSIGSVTAGLSYRFGQ--SAPVV------EPKVVAKTFALNSDVTFAFGKANLRPEAQ 243

Qy  234  QALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKI 293
            |:  | |  :::|| |||||||||||| :|:||||:||:|::::
Db  244  NVLDGIYGEIAQL--KSVQVDVAGYTDRIGSEAANLKLSQRRADTVANYLVSKGVAQEVI 301

Qy  294  SARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG 334
            |||  ||||| ||  |||:||| |||||||||||
Db  302  SSTGYGEANPVTGAKCDTVKGRKALIACLADDRRVEISVKG 342
```

RESULT 10
US-10-432-056-2

```
;   Sequence 2, Application US/10432056
;   Publication No. US20040014661A1
;   GENERAL INFORMATION:
;   APPLICANT:  GOETSCH, Liliane
;   APPLICANT:  HAEUW, Jean-Francois
;   APPLICANT:  ROBERT, Alain
;   TITLE OF INVENTION:  PERIPLASMIC DOMAIN OF AN ENTEROBACTERIUM OMP PROTEIN
;   TITLE OF INVENTION:  AND ITS USE AS CARRIER OR ADJUVANT
;   FILE REFERENCE: 344 667 - US
;   CURRENT APPLICATION NUMBER: US/10/432,056
;   CURRENT FILING DATE: 2003-05-16
;   PRIOR APPLICATION NUMBER: PCT/FR 01/03 596
;   PRIOR FILING DATE: 2001-11-16
;   PRIOR APPLICATION NUMBER: FR 00/14 909
;   PRIOR FILING DATE: 2000-11-17
;   NUMBER OF SEQ ID NOS: 6
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 2
;   LENGTH: 137
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-10-432-056-2

Query Match       37.7%;   Score 687;  DB 15;  Length 137;
Best Local Similarity  100.0%;  Pred. No. 1.8e-55;
Matches   135;  Conservative   0;  Mismatches   0;  Indels  0;  Gaps  0;

Qy  208  VATKHFTLKSDVLFNFKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAY  267
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  VATKHFTLKSDVLFNFKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAY   60

Qy  268  NQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRR 327
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  NQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRR 120

Qy  328  VEIEVKGYKEVVTQP 342
            |||||||||||||||
Db  121  VEIEVKGYKEVVTQP 135
```

RESULT 11
US-09-905-176-22

```
;   Sequence 22, Application US/09905176
;   Patent No. US20020150906A1
```

; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: Debe, Derek A.
; TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FROM P
; TITLE OF INVENTION: PROTEIN SEQUENCE
; FILE REFERENCE: 265/297
; CURRENT APPLICATION NUMBER: US/09/905,176
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/218,016
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-905-176-22
; LENGTH: 153

Query Match　　37.7%;　　Score 687; DB 9; Length 153;
Best Local Similarity　79.1%;　Pred. No. 2.1e-55;
Matches 129; Conservative 10; Mismatches 14; Indels 10; Gaps 2;

Qy　 27 YHDTGFYGNGPQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMGYDWLGRMAYKGSVDNGAF 86
        ||||| ||||||||||||||||||| :::|||||||:||:|:|
Db　 1 YHDTGLI----NNNGPTHENKLGAGAFGGYQVNPYVGFEMGYDWLGRMPYKGSVENGAY 55

Qy　 87 KAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGV 146
        ||||||||||||||||||||||||||||||:: ||||||||||||||||
Db　 56 KAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADTYSN------VYGKNHDTGVSPVFAGGV 110

Qy　147 EWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRF 189
        :|:|: :||:|||||||||||||| |||||||||||||||
Db　111 EYAITPEIATRLEYQWTNNIGDAHTIGTRPDNGMLSLGVSYRF 153

RESULT 12
US-10-336-840-3
; Sequence 3, Application US/10336840
; Publication No. US2003021945A1
; GENERAL INFORMATION:
; APPLICANT: TERRY, TAMSIN DEBORAH
; APPLICANT: TSENG, HSING-JU
; APPLICANT: HOBB, RHONDA IVY
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: DOWNES, JOHN
; TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
; FILE REFERENCE: 37955-0007
; CURRENT APPLICATION NUMBER: US/10/336,840
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: PCT/AU01/00822
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Haemophilus paragallinarum
US-10-336-840-3

Query Match　　37.6%;　　Score 685; DB 14; Length 344;
Best Local Similarity　44.9%;　Pred. No. 1e-54;
Matches 153; Conservative 46; Mismatches 124; Indels 18; Gaps 7;

Qy　 1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPQNNNGPTRNDQLGAGAFGGYQVN 60
        |||||: ||||:|||::|| :|| ::|| |||||||||||||||||||||||||||||::|
Db　 13 LTAASVAQAAPQANTFYAGAKAGWASFHDGLNQFENSQNADGTLRN-SVTYGVFGGYQIT 71

Qy　 61 PYLGFEMGYDWLGR-MAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRAD 119
        |||:||||||||:| ||:|:::: |:|:|||||||||||||||||:| :|::|
Db　 72 DNFAVELGYDDFGRAKSRQGGETVIKHTNHGAHLSLKASYPVLEGLDVYARVGAALIRSD 131

Qy　120 SKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR--DIATRLEYQWVNNIG----DAGTVG 173
        || | |||||||||||||||||||||||: | :| ||||||||| :|

; APPLICANT: TERRY, TAMSIN DEBORAH
; APPLICANT: TSENG, HSING-JU
; APPLICANT: HOBB, RHONDA IVY
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: DOWNES, JOHN
; TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
; FILE REFERENCE: 37955-0007
; CURRENT APPLICATION NUMBER: US/10/336,840
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: PCT/AU01/00822
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Haemophilus paragallinarum
US-10-336-840-2

Query Match   37.5%; Score 684; DB 14; Length 344;
Best Local Similarity   44.6%; Pred. No. 1.2e-54;
Matches 152; Conservative 46; Mismatches 125; Indels 18; Gaps 7;

Qy   1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVN 60
            :|:.|||.|||:|.||:.:|| ||:|.|||||||||:.|||||:
Db   13 LTAASVAQAAPQANTFYAGAKAGWASFHDGLNQFENSQNAYGTLRN-SVTYGVFGGYQIT 71

Qy   61 PYLGFEMGYDWLGRMAYKGSVDN-GAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRAD 119
            |:||| || |:   : :||||.||.||||||:|||.:|:
Db   72 DNFAVELGYDDFGRAKLRQDGETVGKHTNHGAHLSLKASYPVLEGLDVYARVGAALIRSD 131

Qy 120 SKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTR--DIATRLEYQWVNNIG----DAGTVG 173
            |||:|.:|||:||||||:.:   :|||||||||||||||
Db 132 YKPTKRAAPNETHEHSLKVSPVFAGGLEYNLPSLPELALRVEYQWVNKVGRWEKDGSRVD 191

Qy 174 TRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSDVLFNKATLKPEGQ 233
            .:.:|:||:||||:   .|:   .|.| ----EPKVVAKTFALNSDVTFAFGKANLRPEAQ 243
Db 192 YTPSIGSVTAGLSYRFGQ--SAPVV------EPKVVAKTFALNSDVTFAFGKANLRPEAQ 243

Qy 234 QALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKI 293
            ||:.:::.:||||||||||||||||:.|||.:::.||||:|||:|:|||:||||
Db 244 NVLDGIYGEIAQL--KSVQVDLAGYTDRIGSEAANLKLSQRRADTVANYLVSKGVAQEVI 301

Qy 294 SARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKG 334
            |:.:.||||||| ||||||||||||||||
Db 302 SSTGYGEANPVTGAKCDAVKGRKALIACLADDRRVEISVKG 342

Search completed: January 19, 2005, 18:31:06
Job time : 531 secs

RESULT 15
US-10-336-840-4
; Sequence 4, Application US/10336840
; Publication No. US20030219454A1
; GENERAL INFORMATION:
; APPLICANT: TERRY, TAMSIN DEBORAH
; APPLICANT: TSENG, HSING-JU
; APPLICANT: HOBB, RHONDA IVY
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: DOWNES, JOHN
; TITLE OF INVENTION: HAEMAGGLUTININ ANTIGEN
; FILE REFERENCE: 37955-0007
; CURRENT APPLICATION NUMBER: US/10/336,840
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: PCT/AU01/00822
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Haemophilus paragallinarum
US-10-336-840-4

Query Match   37.5%; Score 684; DB 14; Length 344;

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      January 19, 2005, 18:02:20 ; Search time 24 Seconds
                                          (without alignments)
                                          950.558 Million cell updates/sec

Title:          US-09-913-772-2
Perfect score:  1823
Sequence:       1 MKAIFVLNAAPKDNTWYAGG.........DRRVEIEVKGYKEVVTQPAG 344.

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters:          478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued Patents AA:*
      1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
      2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
      3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
      4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
      5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
      6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match Length	Match Length	DB	ID	Description
1	1782	97.8	385	4	US-09-489-039A-7451	Sequence 7451, Ap
2	1781	97.7	335	3	US-08-836-500A-2	Sequence 2, Appli
3	1781	97.7	335	3	US-09-679-750-2	Sequence 2, Appli
4	1279	70.2	384	4	US-09-543-681A-7922	Sequence 7922, Ap
5	984	54.0	179	3	US-09-679-750-4	Sequence 4, Appli
6	984	54.0	179	3	US-09-809-665A-151	Sequence 151, App
7	671	36.8	364	4	US-09-418-980-8	Sequence 8, Appli
8	667	36.6	364	4	US-08-457-997B-2	Sequence 2, Appli
9	662.5	36.3	359	4	US-08-467-722A-2	Sequence 2, Appli
10	662.5	36.3	359	3	US-09-451-184-2	Sequence 2, Appli
11	662.5	36.3	359	4	US-09-809-665A-153	Sequence 153, App
12	646.5	35.5	369	4	US-09-418-980-10	Sequence 10, Appl
13	639.5	35.1	369	1	US-09-210-394-1	Sequence 1, Appli
14	586.5	32.2	338	3	US-08-836-500A-6	Sequence 6, Appli
15	379	20.8	72	3	US-09-679-750-6	Sequence 6, Appli
16	379	20.8	72	4	US-08-836-500A-8	Sequence 8, Appli
17	281	15.4	53	3	US-09-679-750-8	Sequence 8, Appli
18	281	15.4	53	4	US-08-457-997B-2	Sequence 8, Appli
19	242.5	13.3	379	4	US-09-328-352-5219	Sequence 5219, Ap
20	231.5	12.7	351	4	US-09-252-991A-30094	Sequence 30094, A
21	224.5	12.3	349	4	US-09-573-630A-2	Sequence 2, Appli
22	199	10.9	235	4	US-09-252-991A-19595	Sequence 19595, A
23	193.5	10.6	259	4	US-09-328-352-5775	Sequence 5775, Ap
24	184.5	10.1	566	4	US-09-489-039A-14179	Sequence 14179, A
25	182	10.0	220	4	US-09-489-039A-11012	Sequence 11012, A
26	182	10.0	226	2	US-08-572-447C-15	Sequence 15, Appl
27	182	10.0	226	3	US-09-267-747-15	Sequence 15, Appl
28	181	9.9	181	2	US-08-572-447C-11	Sequence 11, Appl
29	181	9.9	181	2	US-09-267-747-11	Sequence 11, Appl
30	176.5	9.7	266	3	US-09-252-991A-18046	Sequence 18046, A
31	175	9.6	633	4	US-09-489-039A-13729	Sequence 13729, A
32	175	9.6	753	4	US-09-252-991A-17612	Sequence 17612, A
33	170	9.3	472	4	US-09-328-352-6870	Sequence 6870, Ap
34	169.5	9.3	272	4	US-09-252-991A-29681	Sequence 29681, Ap
35	168.5	9.2	214	3	US-08-572-447C-13	Sequence 13, Appl
36	168.5	9.1	214	3	US-09-267-747-13	Sequence 13, Appl
37	165.5	9.1	417	4	US-09-252-991A-26245	Sequence 26245, A
38	163.5	9.0	169	4	US-09-252-991A-19373	Sequence 19373, A
39	161	8.8	278	4	US-09-328-352-5824	Sequence 5824, Ap
40	158	8.7	257	4	US-09-328-352-6621	Sequence 6621, Ap
41	149	8.2	231	4	US-09-540-236-3827	Sequence 3827, Ap
42	146.5	8.0	172	4	US-09-674-779B-6	Sequence 6, Appli
43	143.5	7.9	172	4	US-09-674-779B-8	Sequence 8, Appli
44	142.5	7.8	152	4	US-09-540-236-3190	Sequence 3190, Ap
45	142.5	7.8	172	4	US-09-674-779B-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7451
; Sequence 7451, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7451
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7451

Query Match           97.8%;  Score 1782;  DB 4;  Length 385;
Best Local Similarity 99.4%;  Pred. No. 5e-169;
Matches 336;  Mismatches  2;  Conservative  0;  Indels  0;  Gaps  0;

Qy     6 VLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFCGYQVNPYLGF 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    47 VAQAAPKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFCGYQVNPYLGF 106

Qy    66 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   107 EMGYDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYA 166

Qy   126 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   167 STGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGV 226

Qy   186 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   227 SYRFGQEDAAPVVAPAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSN 286

Qy   246 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRRAQSVVDYLVAKGIPAGKISARGMGESNPVT 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   287 MDPKDGSAVVLGYTDRIGSEAYNQQLSEKRRAQSVVDYLVAKGIPAGKISARGMGESNPVT 346

Qy   306 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPA 343
          ||||||||||||||||||||||||||||||||||||||
Db   347 GNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPA 384

RESULT 2
```

US-08-836-500A-2
; Sequence 2, Application US/08836500A
; Patent No. 6197929
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   APPLICANT: Baussant, Thierry
;   APPLICANT: Haeuw, Jean-Francois
;   APPLICANT: Nguyen Ngoc, Thien
;   TITLE OF INVENTION: Carrier Protein Having an Adjuvant
;   TITLE OF INVENTION: Effect, Immunogenic Complex Containing It, Process for
;   TITLE OF INVENTION: Their Preparation, Nucleotide Sequence and Vaccines
; Patent No. 6197929
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
;   STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
;   STREET: 4700
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: U.S.A.
;   ZIP: 60601
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/836,500A
;   FILING DATE:
;   CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
;   NAME: Katz, Martin L.
;   REGISTRATION NUMBER: 25,011
;   REFERENCE/DOCKET NUMBER: PIE1514P0180US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-616-5400
;   TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 335 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-836-500A-2

Query Match   97.7%;   Score 1781;   DB 3;   Length 335;
Best Local Similarity   100.0%;   Pred. No. 5.1e-169;
Matches 335;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy  10 APKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGFEMGY 69
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 APKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGFEMGY 60

Qy  70 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 129
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 120

Qy 130 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRF 189
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRF 180

Qy 190 GQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPK 249
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPK 240

Qy 250 DGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTC 309
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 DGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTC 300

Qy 310 DNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 344
          |||||||||||||||||||||||||||||||
Db 301 DNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 335
```

RESULT 3
US-09-679-750-2
; Sequence 2, Application US/09679750
; Patent No. 6780420
; GENERAL INFORMATION:
;   APPLICANT: Binz, Hans
;   Baussant, Thierry
;   Haeuw, Jean-Francois
;   Nguyen Ngoc, Thien
;   TITLE OF INVENTION: Carrier Protein Having an Adjuvant
;   Effect, Immunogenic Complex Containing It, Process for
;   Their Preparation, Nucleotide Sequence and Vaccines
; Patent No. 6780420
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
;   STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
;   4700
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: U.S.A.
;   ZIP: 60601
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/679,750
;   FILING DATE: 08-Oct-2000
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/836,500
;   FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;   NAME: Katz, Martin L.
;   REGISTRATION NUMBER: 25,011
;   REFERENCE/DOCKET NUMBER: PIE1514P0180US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-616-5400
;   TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 335 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-679-750-2

Query Match   97.7%;   Score 1781;   DB 4;   Length 335;
Best Local Similarity   100.0%;   Pred. No. 5.1e-169;
Matches 335;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy  10 APKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGFEMGY 69
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 APKDNTWYAGGKLGWSQYHDTGFYGNGPTRNDQLGAGAFGGYQVNPYLGFEMGY 60

Qy  70 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 129
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 120

Qy 130 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRF 189
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYRF 180

Qy 190 GQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPK 249
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GQEDAAPVVAPAPAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPK 240

Qy 250 DGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTC 309
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   241 DGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTC 300
```

```
Qy   310 DNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 344
             |||||||||||||||||||||||||||||||||||
Db   301 DNVKARAALIDCLAPDRRVEIEVKGYKEVVTQPAG 335
```

RESULT 4
US-09-543-681A-7922
; Sequence 7922, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7922
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7922

Query Match   70.2%; Score 1279; DB 4; Length 384;
Best Local Similarity 71.5%; Pred. No. 6.7e-119;
Matches 243; Conservative 32; Mismatches 59; Indels 6; Gaps 3;

```
Qy     9 AAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMG  68
           ||||||||||||||||||||||| | :|:||||||||||:|||||||||||||||||
Db    43 AAPKDNTWTGGKLGWSQVQSTGNNWDGVNIGNGSTHKDQIGAGAFAGYQYNQYLGFEIG 102

Qy    69 YDWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKG--NYAS 126
           ||||||||||||||||||||:|||:|| |:||||||||||||||||||||||    ||:|
Db   103 YDWLGRMAYKGSYNNGAFKAQGIQLTTKLSYPVMDDLDVYTRLGGMVWRADSTATINATS 162

Qy   127 TGVSR--SEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLG 184
           ||| :    |||||||||||:|||||||||||||||||||||:|| ||||:| |:|:||
Db   163 AGTQKRFSENDTGVSPVFALGTEYAITPNIATRLEYQWINNIGDKGTLNARPDNGMLSVG 222

Qy   185 VSYRFGQEDAAPVVAPAP--APAPEVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQ 242
           || ||   : :|| ||  :||||||:|  |:| ||:|| | |:| :  |||| ||||
Db   223 VAYRFNQETPAPVVEPAPVVAPAPVVENKTFTLRSDVLFNYNKSSLKAEGQEALNGLYNE 282

Qy   243 LSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESN 302
           | |:| |||||||||||||||||||: :||||||:|||||||:||||||||:|: |:||:
Db   283 LANIDPTQGRVVVIGYTDRIGSQNYNLPLSEKRAQSVVDYLVSKGIPANSISAEGRGKEN 342

Qy   303 PVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKEVVTQP 342
           |||||||||:||||||||||||||||||:::||||| || |
Db   343 PVTGNTCDNIKARAALIDCLAPDRRVEIEIQGTTEVVVQP 382
```

RESULT 5
US-08-836-500A-4
; Sequence 4, Application US/08836500A
; Patent No. 6197929
; GENERAL INFORMATION:
; APPLICANT: Binz, Hans
; APPLICANT: Baussant, Thierry
; APPLICANT: Haeuw, Jean-Francois
; APPLICANT: Nguyen Ngoc, Thien
; TITLE OF INVENTION: Carrier Protein Having an Adjuvant
; TITLE OF INVENTION: Effect, Immunogenic Complex Containing It, Process for
; TITLE OF INVENTION: Their Preparation, Nucleotide Sequence and Vaccines
Patent No. 6197929
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, 2 Prudential Plaza, Suite

; STREET: 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,500A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: PIE1514P0180US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-836-500A-4

Query Match   54.0%; Score 984; DB 3; Length 179;
Best Local Similarity 100.0%; Pred. No. 5.3e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    10 APKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMGY  69
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 APKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMGY  60

Qy    70 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 129
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV 120

Qy   130 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYR 188
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYR 179
```

RESULT 6
US-09-679-750-4
; Sequence 4, Application US/09679750
; Patent No. 6780420
; GENERAL INFORMATION:
; APPLICANT: Binz, Hans
; Baussant, Thierry
; Haeuw, Jean-Francois
; Nguyen Ngoc, Thien
; TITLE OF INVENTION: Carrier Protein Having an Adjuvant
; Effect, Immunogenic Complex Containing It, Process for
; Their Preparation, Nucleotide Sequence and Vaccines

Patent No. 6780420
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, 2 Prudential Plaza, Suite
; 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/679,750
  FILING DATE: 08-Oct-2000
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/836,500
  FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Katz, Martin L.
  REGISTRATION NUMBER: 25,011
  REFERENCE/DOCKET NUMBER: PIE1514P0180US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 312-616-5400
  TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 179 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
  SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-679-750-4

Query Match        54.0%;  Score 984;  DB 4;  Length 179;
Best Local Similarity   100.0%;
Matches  179;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      10  APKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMGY  69
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  APKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGAGAFGGYQVNPYLGFEMGY  60

Qy      70  DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV  129
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  DWLGRMAYKGSVDNGAFKAQGVQLTAKLGYPITDDLDIYTRLGGMVWRADSKGNYASTGV  120

Qy     130  SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYR  188
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRPDNGMLSLGVSYR  179


RESULT 7
US-09-809-665A-151
; Sequence 151, Application US/09809665A
; Patent No. 6790950
; GENERAL INFORMATION:
;  APPLICANT: Lowery E., David, et al.
;  TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
;  FILE REFERENCE: 28341/00435
;  CURRENT APPLICATION NUMBER: US/09/809,665A
;  CURRENT FILING DATE: 2001-03-15
;  PRIOR APPLICATION NUMBER: 60/153,453
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: 60/128,689
;  PRIOR FILING DATE: 1999-04-09
;  PRIOR APPLICATION NUMBER: 09/545,199
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 197
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 151
;  LENGTH: 364
;  TYPE: PRT
;  ORGANISM: Actinobacillus pleuropneumoniae
US-09-809-665A-151

Query Match        36.8%;  Score 671;  DB 4;  Length 364;
Best Local Similarity   42.4%;  Pred. No. 2.4e-58;
Matches  153;  Conservative  54;  Mismatches 122;  Indels 32;  Gaps 10;

Qy       1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTR----NDQLGAGAFGG  56
            : |||||||||||||||  |  |:::  |||||||||   :  |: ||||
Db       1  LSAAAVAQAAPQQNTFYAGAKAGWASFHDGIEQLDSAKNTDRGTKYGINRNSVTYGVFGG  70

Qy      57  YQV--NPYLGF--EMGYDWLGRM----AYKGSVDNGAFK--AQGVQLTAKLGYPITDDLD  106
```

```
Db      71  YQILNQDKLGLAAELGYDYFGRVRGSEKPNGKADKKTFRHAAHGATIALKPSYEVLPDLD  130
            ||:  ||||:|||||||||||||||||||||| ||:||
Qy     107  IYTRLG-GMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN  165
            :|::|  |||||||||||||||||:  :||||||||||||||||||||||||||||
Db     131  VYGKVGIALVNNTYKTFNAAQEKVKTRRFQS--SLILGAGVEYAILPELAARVEYQWLNN  188
            :|:|||  ||||||||||||| |||||||
Qy     166  IGDA-------GTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKGHFTLKS  217
            ||||       |||||||||| ||||||||  |||  ||  ||
Db     189  AGKASYSTLNRMGATDYRSDISSVSAGLSYRFGQ-GAPVAAPA------VETKNFAFSS  241
            ||||||||||||:||||||||||:||||||||| ||||||  |||||||||||
Qy     218  DVLFNFNKCATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQ  277
            ||||||||:|| ||| :|| |||:||:|||||||||::|||::|||||:||||:||||
Db     242  DVLFAFGKSNLKPAAATALDAMQTEINNAGLSNAAIQVNGYTDRIGKEASNLKLSQRRAE  301
            |||||||||| ||||||| ||||:|||||||| ||||||| |||||
Qy     278  SVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKE  337
            ||||||||||||||||||||||||||||||||||||||||||||| |||||||||:
Db     302  TVANYIVSKGAPAANVTAVGYGEANPVTGATCDKVKGRKALIACLAPDRRVEVQVQGTKE  361

Qy     338  V 338
            |
Db     362  V 362


RESULT 8
US-09-418-980-8
; Sequence 8, Application US/09418980
; Patent No. 6713071
; GENERAL INFORMATION:
;  APPLICANT: Campos, Manuel
;  APPLICANT: Baarsch, Mary Jo
;  APPLICANT: Rosey, Everett
;  APPLICANT: Ankenbauer, Robert
;  APPLICANT: Warren-Stewart, Lynn
;  APPLICANT: Suiter, Brian
;  APPLICANT: Keach, Robin
;  TITLE OF INVENTION: NOVEL PROTEINS FROM ACTINOBACILLUS PLEUROPNEUMONIAE
;  FILE REFERENCE: PC9854A
;  CURRENT APPLICATION NUMBER: US/09/418,980
;  CURRENT FILING DATE: 1999-10-14
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 8
;  LENGTH: 364
;  TYPE: PRT
;  ORGANISM: Actinobacillus pleuropneumoniae
US-09-418-980-8

Query Match        36.8%;  Score 667;  DB 4;  Length 364;
Best Local Similarity   42.1%;  Pred. No. 6.1e-58;
Matches  152;  Conservative  54;  Mismatches 123;  Indels 32;  Gaps 10;

Qy       1  MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYGNGFQNNNGPTR----NDQLGAGAFGG  56
            : |||||||||||||||  |  |:::  |||||||||   :  |: ||||
Db      11  LSAAAVAQAAPQQNTFYAGAKAGWASFHDGIEQLDSAKNTDRGTKYGINRNSVTYGVFGG  70

Qy      57  YQV--NPYLGF--EMGYDWLGRM----AYKGSVDNGAFK--AQGVQLTAKLGYPITDDLD  106

Qy      71  YQILNQDKLGLAAELGYDYFGRVRGSEKPNGKADKKTFRHAAHGATIALKPSYEVLPDLD  130

Qy     107  IYTRLG-GMVWRADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN  165

Qy     131  VYGKVGIALVNNTYKTFNAAQEKVKTRRFQS--SLILGAGVEYAILPELAARVEYQWLNN  188

Qy     166  IGDA-------GTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKGHFTLKS  217

Db     189  AGKASYSTLNRMGATDYRSDISSVSAGLSYRFGQ-GAPVAAPA------VETKNFAFSS  241

Qy     218  DVLFNFNKCATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQ  277

Db     242  DVLFAFGKSNLKPAAATALDAMQTEINNAGLSNAAIQVNGYTDRIGKEASNLKLSQRRAE  301

Qy     278  SVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYKE  337
```

Db   302 VANYFVAKGVAADAISATGYGEANPVTGATCDQVKGRKALIACLAPDRRVEIAVNGTK 359

RESULT 10
US-08-467-722A-2
; Sequence 2, Application US/08467722A
; Patent No. 6030626
; GENERAL INFORMATION:
;   APPLICANT: Kolattukudy, P. E.
;   TITLE OF INVENTION: Otitis Media Vaccine
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calfee, Halter and Griswold
;   STREET: Suite 1800 800 Superior Avenue
;   CITY: Cleveland
;   STATE: Ohio
;   COUNTRY: U.S.A>
;   ZIP: 44114-2688
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/467,722A
;   FILING DATE:
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Golrick, Mary E.
;   REGISTRATION NUMBER: 34,829
;   REFERENCE/DOCKET NUMBER: 22727/00102
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (216) 622-8458
;   TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 359 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-467-722A-2

Query Match   36.3%; Score 662.5; DB 3; Length 359;
Best Local Similarity 43.3%; Pred. No. 1.7e-57;
Matches 155; Conservative 49; Mismatches 121; Indels 33; Gaps 9;

Qy   1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHD----TGFYGNGFQNNNGPTRNDQLGAGAFGG 56
Db   13 LAAASVAQAAPQENTFYAGVKAGQGSFHDGINNNGAIKKGLSSSNYGYRRNTFTYGVFGG 72

Qy   57 YQV----NPYLGFEMGYDWLGRMAYKGSVDNGAFKAQ---GVQLTAKLGYPITDDLDIY 108
Db   73 YQILNQDNFGLAAELGYDDFGRAKLR---EAGKPKAKHTNHGAYLSLKGSYEVLDGLDVY 129

Qy 109 TRLGGMVWRADSKGNYASTGV---SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN 165
Db 130 GKAGVALVRSDYKFYEDANGTRDHKKGRHTARASGLFAVGAEYAVLPELAVRLEYQWLTR 189

Qy 166 IG------DAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD 218
Db 190 VGKYRPQDKPNTAINYNPWIGCINAGISYRFGQGE-APVVA-----APEMVSKTFSLNSD 243

Qy 219 VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQS 278
Db 244 VTFAFGKANLKPQAQATLDSVYGEISQV--KSRKVAVAGYTNRIGSDAFNVKLSQERADS 301

Qy 279 VVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYK 336
Db 302 VANYFVAKGVAADAISATGYGEANPVTGATCDQVKGRKALIACLAPDRRVEIAVNGTK 359

RESULT 11
US-09-451-184-2

Db   302 TVANYIVSKGAPAANVTAVGYGEANPVTGATCDKVKGRKALIACLAPDRRVEVQVQGTKE 361

Qy   338 V 338
Db   362 V 362

RESULT 9
US-08-457-997B-2
; Sequence 2, Application US/08457997B
; Patent No. 5766608
; GENERAL INFORMATION:
;   APPLICANT: Kolattukudy, P. E.
;   TITLE OF INVENTION: Otitis Media Vaccine
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Calfee, Halter and Griswold
;   STREET: Suite 1800 800 Superior Avenue
;   CITY: Cleveland
;   STATE: Ohio
;   COUNTRY: U.S.A.
;   ZIP: 44114-2688
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/457,997B
;   FILING DATE:
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Golrick, Mary E.
;   REGISTRATION NUMBER: 34,829
;   REFERENCE/DOCKET NUMBER: 22727/00102
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (216) 622-8458
;   TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 359 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-457-997B-2

Query Match   36.3%; Score 662.5; DB 1; Length 359;
Best Local Similarity 43.3%; Pred. No. 1.7e-57;
Matches 155; Conservative 49; Mismatches 121; Indels 33; Gaps 9;

Qy   1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHD----TGFYGNGFQNNNGPTRNDQLGAGAFGG 56
Db   13 LAAASVAQAAPQENTFYAGVKAGQGSFHDGINNNGAIKKGLSSSNYGYRRNTFTYGVFGG 72

Qy   57 YQV----NPYLGFEMGYDWLGRMAYKGSVDNGAFKAQ---GVQLTAKLGYPITDDLDIY 108
Db   73 YQILNQDNFGLAAELGYDDFGRAKLR---EAGKPKAKHTNHGAYLSLKGSYEVLDGLDVY 129

Qy 109 TRLGGMVWRADSKGNYASTGV---SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN 165
Db 130 GKAGVALVRSDYKFYEDANGTRDHKKGRHTARASGLFAVGAEYAVLPELAVRLEYQWLTR 189

Qy 166 IG------DAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD 218
Db 190 VGKYRPQDKPNTAINYNPWIGCINAGISYRFGQGE-APVVA-----APEMVSKTFSLNSD 243

Qy 219 VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAYNQQLSEKRAQS 278
Db 244 VTFAFGKANLKPQAQATLDSVYGEISQV--KSRKVAVAGYTNRIGSDAFNVKLSQERADS 301

Qy 279 VVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYK 336
Qy

; Sequence 2, Application US/09451184
; Patent No. 6562349
; GENERAL INFORMATION:
; APPLICANT: Kolattukudy, P. E.
; TITLE OF INVENTION: Otitis Media Vaccine
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 1400 McDonald Investment Center,
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A>
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/451,184
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Docherty, Pamela A.
; REGISTRATION NUMBER: 40,591
; REFERENCE/DOCKET NUMBER: 24547/04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8416
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-451-184-2

Query Match   36.3%;   Score 662.5; DB 4;   Length 359;
Best Local Similarity 43.3%;   Pred. No. 1.7e-57;
Matches 155; Conservative 49; Mismatches 121; Indels 33; Gaps 9;

Qy   1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHD----TGFYGNGFQNNNGPTRNDQLGAGAFGG 56
          || :||:||||||| :||||| : :|| | |||||
Db   13 LSAAAVAQAAPQQNTFYAGAVKGAGQGSFHDGINNNGAIKKGLSSSNYGYRRNTFTYGVFGG 72

Qy   57 YQV---NPYLGFEMGYDWLGRMAYKGSVDNGAFKAQ----GVQLTAKLGYPITDDLDIY 108
          :|| :|||:||||||||||||||| | :|||||||||||||: |:|:
Db   73 YQILNQDNFGLAAELGYDDFGRAKLR---EAGKPKAKHTNHGAYLSLKGSYEVLDGLDVY 129

Qy 109 TRLGGMVWRADSKGNYASTGV---SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN 165
          ||||||||||||||||||| :|| |||||||||||||||||||: :|:|||:|
Db 130 GKAGAVLVRSDYKFYEDANGTRDRHKKGRHTARASGLFAVGAEYAVLPELAVRLEYQWLTR 189

Qy 166 IG------DAGTVGTRPDNGMLSLGVSYRFGQEDAAPVVAPAPAPEVATKHFTLKSD 218
          :|:   : :||||||||||||||||||||: :| :|:|
Db 190 VGKYRPQDKPNTAINYNPWIGCINAGISYRFGQGE-APVVA-----APEMVSKTFSLNSD 243

Qy 219 VLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAVNQQLSEKRAQS 278
          || |||||||||| ||| |||||||||||||||||||||||||| :|||||||||
Db 244 VTFAFGKANLKPQAQATLDSVYGEISQV--KSRKVAVAGVTNRIGSDAFNVKLSQERADS 301

Qy 279 VVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEVKGYK 336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||:| |
Db 302 VANYFVAKGVAADAISATGYGEANPVTGATCDQVKGRKALIACLAPDRRVEIAVNGTK 359

RESULT 12
US-09-809-665A-153
; Sequence 153, Application US/09809665A
; Patent No. 6790950
; GENERAL INFORMATION:
; APPLICANT: Lowery E., David, et al.

; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions
; FILE REFERENCE: 28341/00435
; CURRENT APPLICATION NUMBER: US/09/809,665A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/153,453
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/128,689
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 09/545,199
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 153
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumoniae
US-09-809-665A-153

Query Match   35.5%;   Score 646.5; DB 4;   Length 369;
Best Local Similarity 41.1%;   Pred. No. 6.9e-56;
Matches 153; Conservative 60; Mismatches 110; Indels 49; Gaps 12;

Qy   1 MKAIFVLNAAPKDNTWYAGGKLGWSQYHDTGFYG-NGFQNNNGPTRNDQ-------L 49
          |:||| | ||||||||||||||| :|: | :|
Db   11 LSAAAVAQAAPQQNTFYAGAKVGQSSFH---HGVNQLKSGHDDRYNDKTRKYGINRNSV 66

Qy   50 GAGAFGGYQV---NPYLGFEMGYDWLGRMAYKGSVDNGAFK------AQGVQLTAKLGYP 100
          |||| ||||| :|||:|||||||||| : || | ||||||||||||
Db   67 TYGVFGGYQILNQNNFGLAAELGYDYYGRV-RGNVDE--FRTVKHSAHGLNLALKPSYE 122

Qy 101 ITDDLDIYTRLGGMVWRADSKGNYASTGVSRSE--HDTGVSPVFAGGVEWAVTRDIATR 157
          :|: :|:||||||||||||||||||||| ::|:|||||| |::::||||:||
Db 123 VLPDLDVYGKVGIAVRNDYK-KYGAENTNESTTKFHKLLKASTILGAGVEYAILPELAAR 181

Qy 158 LEYQWVNNIGDAGTV-------GTR--------PDNGMLSLGVSYRFGQEDAAPVVAPAPAPAP 206
          |||||:   : || ||| :|:|||||||||||||: :||||||
Db 182 VEYQYLNKAGNLNKALVRSGTQDVDFQYAPDIHSVTAGLSYRFGQGAVAPVV------EP 235

Qy 207 EVATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEA 266
          ||||||||||| || |||:|||||||||||||||||||||||||||||||||||||||||
Db 236 EVTKNFAFSSDVLFDFGKSSLKPAAATALDAANTEIANLGLATPAIQVNGYTDRIGKEA 295

Qy 267 YNQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDR 326
          :|::|||||| ::|:|| |||||||||||||||||||:||||||:||::|||:||||||
Db 296 SNLKLSQRRAETVANVIVSKGQNPANVTAVGYGEANPVTGATCDAVKGRKALIACLAPDR 355

Qy 327 RVEIEVKGYKEV 338
          |||:::|| ||
Db 356 RVEVQVGAKNV 367

RESULT 13
US-09-418-980-10
; Sequence 10, Application US/09418980
; Patent No. 6713071
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Baarsch, Mary Jo
; APPLICANT: Rosey, Everett
; APPLICANT: Ankenbauer, Robert
; APPLICANT: Warren-Stewart, Lynn
; APPLICANT: Suiter, Brian
; APPLICANT: Keach, Robin
; TITLE OF INVENTION: NOVEL PROTEINS FROM ACTINOBACILLUS PLEUROPNEUMONIAE
; FILE REFERENCE: PC9854A
; CURRENT APPLICATION NUMBER: US/09/418,980
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumoniae

US-09-418-980-10

```
Query Match          35.1%;  Score 639.5;  DB 4;  Length 369;
Best Local Similarity 40.7%;  Pred. No. 3.4e-55;
Matches 151;  Conservative 60;  Mismatches 113;  Indels 47;  Gaps 11;

Qy    1 MKAIFVLNAAPKDNTWTAGGKLGWSQYHDTGFYG-NGFQNNNGPTRNDQ--------L   49
          :  ::: :|||  ||||:|| :| :|:     ||  :   :: :||    
Db   11 LSAAAVAQAAPQQNTFYAGAKVQGSSFH----HGVNQLKSGHDDRYNDKTRKYGINRNSV  66

Qy   50 GAGAFGGYQV----NPYLGFEMGYDWLGRMAYKGSVDNGAFK-----AQGVQLTAKLGYP 100
          | ||||||:|     ||| ||:|| ||||| ::| ||  :       | ||: ||:|||
Db   67 TYGVFGGYQILNQNNFGLATELGYDYIGRV--RG--NDGEFRAMKHSAHGLNFALKPSYE 122

Qy  101 ITDDLDIYTRLGGMVWRADSKGNYASTGVSRSE--HDTGVSPVFAGGVEWAVTRDIATRL 158
          : |||| : :|| :  : ::||||   : : |   ||| |||: ||||| |  ||
Db  123 VLPDLDVVGKVGVAVVRNDYKSYGAENTNEPTEKFHKLKASTILGAGVEYAILPELAARV 182

Qy  159 EYQWVNNIGDAGTV----GTR------PDNGMLSLGVSYRFGQEDAAPVVAPAPAPAPE 207
          |||::|  | : ||    ||:       || ||||:|:||||| |  ::|  |  :
Db  183 EYQYLNKAGNLNKALVRSGTQDVDFQYAPDIHSVTAGLSYRFGQGAVAPVV-----EPE 236

Qy  208 VATKHFTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAY 267
          | || |: :SDVLF:F K:|LKP    | |    :    |  :  | ||||||||| ||
Db  237 VVTKNFAFRSSDVLFDFGKSSLKPAAATALDAANTEIANLGLATPAIQVNGYTDRIGKEAS 296

Qy  268 NQQLSEKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRR 327
          | :||::|||: ||:|||:|| |||||||||||||||||| ||:||||::|| |||||||
Db  297 NLKLSQRRAETVANYLVSKGQNPANVTAVGYGEANPVTGATCDKVKGRKALIACLAPDRR 356

Qy  328 VEIEVKGYKEV 338
          ||:|: | |:|
Db  357 VEVQVQGAKNV 367
```

; MOLECULE TYPE: protein

US-08-210-394-1

```
Query Match          32.2%;  Score 586.5;  DB 1;  Length 338;
Best Local Similarity 40.1%;  Pred. No. 5.8e-50;
Matches 146;  Conservative 44;  Mismatches 111;  Indels 63;  Gaps 12;

Qy   10 APKDNTWTAGGKLGWSQYHDTGFYGNGFQNNNGPTRNDQLGA--------GAFGG  56
          |:::||: ||.||. | |:.|      |  :||||| :|    :        |:||
Db    1 APQENTFYAGVKAGGQGSFHD------GI-NNNGAIKEDSIDLTLGYGYRRNTFTYGVFGG  53

Qy   57 YQV----NPYLGFEMGYDWLGRMAYKGSVDNGAFKAQ---GVQLTAKLGYPITDLDIY 108
          ||:     |: | | ||||: |||  :  : |         |:|:|| ||::||||:|
Db   54 YQILNQDNFGLAAELGYDNFGRVKFRA---EGKTKAKHTNHGAHiSLKGSYEVLDGLDVY 110

Qy  109 TRLGGMVWRADSKGNYASTGV---SRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNN 165
          .: |  : |:| | :   :      :|    |    :|||:| || :    |||::|:
Db  111 GKAGVALVRSDYKFYEAPNSTRDXKKGTHTARASGLFAVGAEYAVLPELAVRLEYQQLTR 170

Qy  166 IG------DAGTVGTRPDN---------GMLSLGVSYRFGQEDAAPVVAPAPAPEVATKH 212
          :|        :    |||         ||:||| ||:||||: |: :|         ||
Db  171 VGKYRPQDKNAPSINPNTAIHYNPXIGSINAGISYRFGQ-GAAPV---------KT 216

Qy  213 FTLKSDVLFNFNKATLKPEGQQALDQLYTQLSNMDPKDGSAVVLGYTDRIGSEAVNQQLS 272
          |:| ||: : : |.:||||  || |  |   :   : :  | |||||||||| ::|:|:|
Db  217 FSLNLDVTFAFGKANLKPQAQATLDSIYGEMSQV--KSAKVAVAGYTDRIGSDAFNVKLS 274

Qy  273 EKRAQSVVDYLVAKGIPAGKISARGMGESNPVTGNTCDNVKARAALIDCLAPDRRVEIEV 332
          ::|| || :| |||||:  : :|:|: | ||||||| || : | |:||||||||||||:|
Db  275 QERADSVANYFVAKGVAADAISATGYGKANPVTGATXDQVWGRWALIATLAPDRRVEIAV 334

Qy  333 KGYK 336
          .|.|
Db  335 NGTK 338
```

```
;       TELEFAX:  312-616-5460
; INFORMATION FOR SEQ ID NO:  6:
;  SEQUENCE CHARACTERISTICS:
;       LENGTH:  72 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;       MOLECULE TYPE:  protein
US-08-836-500A-6

Query Match      20.8%;   Score 379;  DB 3;  Length 72;
Best Local Similarity  100.0%;  Pred. No. 2.7e-30;
Matches   72;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   117 RADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRP 176
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 RADSKGNYASTGVSRSEHDTGVSPVFAGGVEWAVTRDIATRLEYQWVNNIGDAGTVGTRP  60

Qy   177 DNGMLSLGVSYR 188
         ||||||||||||
Db    61 DNGMLSLGVSYR  72

Search completed: January 19, 2005, 18:17:29
Job time : 26 secs
```